

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 08:49:16 ; Search time 14922 Seconds
(without alignments)
11684.568 Million cell updates/sec

Title: US-10-614-524-1

Perfect score: 3687

Sequence: 1 ttgacttcaaataggaaaa.....tactccttatggaggagaatag 3687

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	100.0	3687	6	AX189649 Sequence
2	3666.2	99.4	4173	6	AR359382 Sequence
3	3666.2	99.4	4173	6	AX098667 Sequence
4	3266.6	88.6	3684	1	AF077326 Bacillus
5	3266.6	88.6	3684	6	I90319 Sequence 7
6	3260.2	88.4	3684	6	AR359400 Sequence
7	3260.2	88.4	3684	6	AX098692 Sequence
8	3136.8	85.1	3687	1	BTCTRY44
9	3135.2	85.0	3687	1	AF368257 Bacillus
10	3133.6	85.0	4074	1	BTCTRY1BA2
11	3133.6	85.0	4074	6	I42103 Sequence 1
12	3133.6	85.0	4798	1	AF363025 Bacillus
13	2687.2	72.9	3934	1	BACCRYIE
14	2687.2	72.9	3934	6	AR068828 Sequence
15	2687.2	72.9	3934	6	I38760 Sequence 3
16	2687.2	72.9	3934	6	I70138 Sequence 3
17	2687.2	72.9	3934	6	AX138580 Sequence
18	2687.2	71.3	3627	6	AX146772 Sequence
19	2687.2	71.3	3627	6	AX146772 Sequence
20	2687.2	70.0	8306	1	BTU70726 Bacillus th

20 2563.2 69.5 3696 1 AY138457
21 2544.4 69.0 3663 6 AX088012 Sequence
22 2541.2 68.9 3663 6 AX088010 Sequence
23 2384.2 64.7 3624 6 AX138578 Sequence
24 2384.2 64.7 3624 6 AX146770 Sequence
25 2271.6 61.6 3842 1 BTMCRYIBC
26 2145 58.2 2549 1 AY570734 Bacillus
27 2004.8 54.4 3651 6 AX138576 Sequence
28 2004.8 54.4 3651 6 AX146768 Sequence
29 2000.4 54.3 4623 1 BTU28801
30 1796.4 48.7 3525 1 BACCRYIBB
31 1793.4 48.6 3522 6 AR085762
32 1793.4 48.6 3522 6 I11609
33 1793.4 48.6 3522 6 I73040
34 1791.8 48.6 3522 6 AR362261
35 1790.2 48.6 3522 6 A27529
36 1662 45.1 4106 1 BACCRYIC
37 1662 45.1 4106 6 I82521
38 1662 45.1 4106 6 I93759
39 1581.6 42.9 3558 6 AR451656
40 1581.6 42.9 3558 6 BD090512
41 1501 40.7 3766 6 BTU31527
42 1495.6 40.6 3624 6 AR029509
43 1495.6 40.6 3624 6 AR098462
44 1495.6 40.6 3624 6 BD194455
45 1495.6 40.6 3624 6 I41422

ALIGNMENTS

RESULT 1
AX189649
LOCUS AX189649 3687 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from Patent WO0147952.
ACCESSION AX189649
VERSION AX189649.1 GI:15143038
KEYWORDS
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
REFERENCE 1
AUTHORS Arnaut,G., Boets,A., Damme,N., Mathieu,E., Vanneste,S. and van
Rie,J
TITLE Insecticidal proteins from Bacillus thuringiensis
JOURNAL Patent: WO 0147952-A 1 05-JUL-2001;
Aventis CropScience N.V. (BE)
FEATURES
source
1. 3687
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1. 3687
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC50778.1"
/db_xref="GI:15143038"
/translation="MTSNKNNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEG
NNINPLVASSTVQTGINIAGLVGFAGIAPSYFVLGMPGRDQWEIFLE
LELDPLNAMPLFAIRNQBPVLMVYQAQANLHLLLRDASLFGSEFGITSGEIOIYIE
RQVQTRDYSYCYEWYNTGLNLTGNTAASWVRYNQFRDLITGLVDLVALFYSYDT
RTYPTNSAQITREYITDAIGTGNVWASMMWNNAPSPSAIETAVIRSRPHLLDFLE
QLATFSSRSWATHTMYRGTIQSRIGGLNTSTHSTNTSINVRISPFSDRD
YWTSEYAGVLIWGYLEPIHGVPTFRFNRPQNTFERGTANYSDPTNSQGLKDS
TELPETTERPNYESYSHRSHIGLISQSRVHVVPVSWTHRSADQNTISGSLTOIP
LVKSPNLSGTSVVGPGTGDIIITNVGSLSMGNFNNTSLQVRVRYVAAASQ
TMURVTVGGSTTFDQGPPTMSANESLTSQSFPAEPFGISAGSGTAGISINNA
GRQTFPHDKIEFIPITATFEAYDLERAQEAENALFTNTNRRRLTDVTHIDQVSN
LVACLSDFCLDEKREKLKVKYAKSLDERNLQDPNFTSINKQDPFISTNEQSNFT
SIHQSEHGWSGSENIITQEGNDVFKENYVTLPGTNECYPTLYQKIGSELKAYTR

YOLRGYIBDSQDLBIYLIRYNKHEFLDVPGTESVPLSVBSPICRGCEPNRCAPHFE WNPDJDCSRDGBEKCAHSHHFLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLE FIEPLKJLAGEALSFKRAEKKWRDREKLUETKRVYVTEAKEADVALPVDQVNRLOA DTNIGMIAADKLVRHIREAYLSLSVPGVNAEIRFEELGRIITASLXDRNVNKN GFNNGLAGKVNKGVHDVQOQSHRSVLVPEWEAEVSQVRVCPGRGVLIRVTAKEG YEGGCVTHIEIENWDELKPNCEBEVYPTDGTCDNDYTAHQGTAVGNSRNAGVEDA YEVDFTASVNYRPTTBEETVYDVRDNHCEYDRGVNYPPLPAGNYMKLELYFPETDK VWIEIGETEGKFIVDSEVLLMEB"									
ORIGIN	Query Match	100.0%;	Score 3687;	DB 6;	Length 3687;				
	Best Local Similarity	100.0%;	Pred. No. 0;						
	Matches 3687;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
QY	1	TTGACTTCAAAATAGGAAATGGAATGAAATATATAATGCTTTATCGATTCGAGCTGTA	60						
DB	1	TTGACTTCAAAATAGGAAATGGAATGAAATATATAATGCTTTATCGATTCGAGCTGTA	60						
QY	61	TCGAATCATTCACACAAATGGATCTATCACGAGTCTGTAATGAGGATTCCTTGTGT	120						
DB	61	TCGAATCATTCACACAAATGGATCTATCACGAGTCTGTAATGAGGATTCCTTGTGT	120						
QY	121	ATAGCGAGGGGNAATATCAATCACTTGTAGCGCATCAACAGTCCAAACGGGTATT	180						
DB	121	ATAGCGAGGGGNAATATCAATCACTTGTAGCGCATCAACAGTCCAAACGGGTATT	180						
QY	181	AACATAGCTGGTAGAATATAGTGTATTAGCGTACCGTTTGTGTCGACAAATAGTACT	240						
DB	181	AACATAGCTGGTAGAATATAGTGTATTAGCGTACCGTTTGTGTCGACAAATAGTACT	240						
QY	241	TTTTATAGTTTTCTTGTGTGTAATATGCGCCCGCGGACAGATCAGTGGGAAATTTTC	300						
DB	241	TTTTATAGTTTTCTTGTGTGTAATATGCGCCCGCGGACAGATCAGTGGGAAATTTTC	300						
QY	301	CTAGAACATGTCGAACCACTTATAATCAACAAATTAACAGAAATGCTAGGAATACGCA	360						
DB	301	CTAGAACATGTCGAACCACTTATAATCAACAAATTAACAGAAATGCTAGGAATACGCA	360						
QY	361	CTTGCTCGAATACAAAGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAT	420						
DB	361	CTTGCTCGAATACAAAGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAT	420						
QY	421	TGGTAGAANAACGGTAGATGATCAAGAACGAGAGTGTCTTTATACCCCAATATAGCC	480						
DB	421	TGGTAGAANAACGGTAGATGATGATCAAGAACGAGAGTGTCTTTATACCCCAATATAGCC	480						
QY	481	TTAGAACTTGATTTCTTAATGGATGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	540						
DB	481	TTAGAACTTGATTTCTTAATGGATGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	540						
QY	541	TTTATTAATGGTATATGCTCAAGCTGCAAAATTTACCTTATTTATGAGAGATGCCTCT	600						
DB	541	TTTATTAATGGTATATGCTCAAGCTGCAAAATTTACCTTATTTATGAGAGATGCCTCT	600						
QY	601	CTTTTTGGTAGGAATTTGGGCTTACATCGCAGGAAATCAACGTTTATGAGCGCCA	660						
DB	601	CTTTTTGGTAGGAATTTGGGCTTACATCGCAGGAAATCAACGTTTATGAGCGCCA	660						
QY	661	GTGGACAAACGAGAGATTTTCCGACTTTTGGTAGAATGTTAATACAGTCTTAAT	720						
DB	661	GTGGACAAACGAGAGATTTTCCGACTTTTGGTAGAATGTTAATACAGTCTTAAT	720						
QY	721	AGCTTGAGAGGGAACAAATCGCCAGTTGGGTGCGTTTATTAATCAATTCGCTAGATCTA	780						
DB	721	AGCTTGAGAGGGAACAAATCGCCAGTTGGGTGCGTTTATTAATCAATTCGCTAGATCTA	780						
QY	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCATTTCCA	840						
DB	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCATTTCCA	840						
QY	841	ATAAATAGAGTGTCTAGTTTAAACAGGGAAATTTTATACAGCGCAATTTGGAGCAACAGG	900						
DB	841	ATAAATAGAGTGTCTAGTTTAAACAGGGAAATTTTATACAGCGCAATTTGGAGCAACAGG	900						

QY	901	GTAAATATGGCAAGTATGAATTGGTATAATAATAATGACCTTCGTTTTCCGTATAGAG	960						
DB	901	GTAAATATGGCAAGTATGAATTGGTATAATAATAATGACCTTCGTTTTCCGTATAGAG	960						
QY	961	ACTGCGGTTATCCGAGCCCGCATCTACTGATTTTCTAGAAACAACTTACAATTTTACG	1020						
DB	961	ACTGCGGTTATCCGAGCCCGCATCTACTGATTTTCTAGAAACAACTTACAATTTTACG	1020						
QY	1021	ACTTCATACGATGGAGTGCTACTAGGCATATGATCTGCGAGGATTAATGGAATCATATG	1080						
DB	1021	ACTTCATACGATGGAGTGCTACTAGGCATATGATCTGCGAGGATTAATGGAATCATATG	1080						
QY	1081	TCTCGGCCAAATAGGAGCGGATTAATAATCACTCAACGATGGTCTCAATATCTTAT	1140						
DB	1081	TCTCGGCCAAATAGGAGCGGATTAATAATCACTCAACGATGGTCTCAATATCTTAT	1140						
QY	1141	AATCCTCTAAGATTTATCACTTCTCTCGAGAGTATATGGAATCATATGCAAGGA	1200						
DB	1141	AATCCTCTAAGATTTATCACTTCTCTCGAGAGTATATGGAATCATATGCAAGGA	1200						
QY	1201	GTGCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTTAAT	1260						
DB	1201	GTGCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTTAAT	1260						
QY	1261	TTTAGGAAACCTCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG	1320						
DB	1261	TTTAGGAAACCTCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG	1320						
QY	1321	TCACCTGGGCTTCAATTTAAAGAGTTTCAAGAACTGAAATTAACAACGAAACCAACAGAACGA	1380						
DB	1321	TCACCTGGGCTTCAATTTAAAGAGTTTCAAGAACTGAAATTAACAACGAAACCAACAGAACGA	1380						
QY	1381	CCAAATATGAATCATATAGTATAGTATCTCACAATAGGGCTCATTTCAATATCTAGT	1440						
DB	1381	CCAAATATGAATCATATAGTATAGTATCTCACAATAGGGCTCATTTCAATATCTAGT	1440						
QY	1441	GTGCATGTACAGTATATTTCTTGGAGCGACCGTAGTGAGATCGTACAAATACCAATAGT	1500						
DB	1441	GTGCATGTACAGTATATTTCTTGGAGCGACCGTAGTGAGATCGTACAAATACCAATAGT	1500						
QY	1501	TCAGATAGCATACACAAATACCAATTTGTTAAATCATTTCAACCTTAAATTCAGGTACCTCT	1560						
DB	1501	TCAGATAGCATACACAAATACCAATTTGTTAAATCATTTCAACCTTAAATTCAGGTACCTCT	1560						
QY	1561	GTAGTCAGTGGCCAGGATTTTACAGGAGGGATATATCCGAACTAACGTTAATGTTAGT	1620						
DB	1561	GTAGTCAGTGGCCAGGATTTTACAGGAGGGATATATCCGAACTAACGTTAATGTTAGT	1620						
QY	1621	GTACTAAGTATGGTCTTAAATTTTAAATTAATACATCATTTACAGCGGTATCGGTGAGAT	1680						
DB	1621	GTACTAAGTATGGTCTTAAATTTTAAATTAATACATCATTTACAGCGGTATCGGTGAGAT	1680						
QY	1681	CGTTATGCTGCTTCAACCAATGGTCTTGGAGGTAATCTGCGAGGGAGTACTACTTTT	1740						
DB	1681	CGTTATGCTGCTTCAACCAATGGTCTTGGAGGTAATCTGCGAGGGAGTACTACTTTT	1740						
QY	1741	GATCAAGGATTTCCCTAGTACTATGAGTCAATAGTCTTTGACATCTCAATCATTTAGA	1800						
DB	1741	GATCAAGGATTTCCCTAGTACTATGAGTCAATAGTCTTTGACATCTCAATCATTTAGA	1800						
QY	1801	TTTGAGGAATTTCTGTTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA	1860						
DB	1801	TTTGAGGAATTTCTGTTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA	1860						
QY	1861	AGTAAATATGCAAGTAGCAAACTTTTACATTTTGAATAAATTTGAATTCATTCATTTACT	1920						
DB	1861	AGTAAATATGCAAGTAGCAAACTTTTACATTTTGAATAAATTTGAATTCATTCATTTACT	1920						
QY	1921	GCAACCTTCGAGAGCAGATACGATTTAGAAAGGGGCAAGGCGGTGAATGCTCTGTTT	1980						
DB	1921	GCAACCTTCGAGAGCAGATACGATTTAGAAAGGGGCAAGGCGGTGAATGCTCTGTTT	1980						

QY 1981 ACTAATACGAATCCAGAGAGATTGAAACAGATGTGACAGATTATCATATTAATGATCAAGTA 2040
Db 1981 ACTAATACGAATCCAGAGAGATTGAAACAGATGTGACAGATTATCATATTAATGATCAAGTA 2040
QY 2041 TCCAAATTTAGTGGCGGTGTTATCCGATCAATCTCTGTTAGATGAAAGAGAGATTACTT 2100
Db 2041 TCCAAATTTAGTGGCGGTGTTATCCGATCAATCTCTGTTAGATGAAAGAGAGATTACTT 2100
QY 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTATGAAAGAACTTACTCCAAGATCCAAAC 2160
Db 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTATGAAAGAACTTACTCCAAGATCCAAAC 2160
QY 2161 TTCACATCCATCAATGAAGAACAGACTTCATATCTAATGAGCAATCGAATTTTACA 2220
Db 2161 TTCACATCCATCAATGAAGAACAGACTTCATATCTAATGAGCAATCGAATTTTACA 2220
QY 2221 TCTATCCATGAACATCTGAAACATGATGGTGGGAGTGGAGATGAGCAATTTACATCCAGAA 2280
Db 2221 TCTATCCATGAACATCTGAAACATGATGGTGGGAGTGGAGATGAGCAATTTACATCCAGAA 2280
QY 2281 GGAATGACGTATTTAAAGAGAAATTAACGTACACTACCGGGGACTTTTAAATGAGTGTAT 2340
Db 2281 GGAATGACGTATTTAAAGAGAAATTAACGTACACTACCGGGGACTTTTAAATGAGTGTAT 2340
QY 2341 CCGAGTATTTATATCAAAAAATAGGAGAGTGGAAATTTAAAGCTTTAFACTCGCTACCAA 2400
Db 2341 CCGAGTATTTATATCAAAAAATAGGAGAGTGGAAATTTAAAGCTTTAFACTCGCTACCAA 2400
QY 2401 TTAGAGGGTATTTGAGATAGTCAAGATTTAGAGATATATTTGATTCGTTTAAATGCG 2460
Db 2401 TTAGAGGGTATTTGAGATAGTCAAGATTTAGAGATATATTTGATTCGTTTAAATGCG 2460
QY 2461 AAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAAGTTGAAAGC 2520
Db 2461 AAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAAGTTGAAAGC 2520
QY 2521 CCAATCGGAAGTGGCGGAAACCGAATCGATGCGGACCAACATTTTGAATGAAATCCCTGAT 2580
Db 2521 CCAATCGGAAGTGGCGGAAACCGAATCGATGCGGACCAACATTTTGAATGAAATCCCTGAT 2580
QY 2581 CTAGATTCCTCTGAGAGATGGGAAATGTTGGGATCATTTCCATCATTTCTCTTTG 2640
Db 2581 CTAGATTCCTCTGAGAGATGGGAAATGTTGGGATCATTTCCATCATTTCTCTTTG 2640
QY 2641 GATATTGATTTGATGACAGACTTGCATGAGATCTAGGCGTGGTGGTGTATTTCAAG 2700
Db 2641 GATATTGATTTGATGACAGACTTGCATGAGATCTAGGCGTGGTGGTGTATTTCAAG 2700
QY 2701 ATTAAGACCGAGGAAGGTCTGCAAGACTAGGAACTTGGAAATTTTAAATGAAAGAAACCA 2760
Db 2701 ATTAAGACCGAGGAAGGTCTGCAAGACTAGGAAATTTTAAATGAAAGAAACCA 2760
QY 2761 TTATTAGAGAGAGACTGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Db 2761 TTATTAGAGAGAGACTGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
QY 2821 GAAAACTTACAAATTTGAAACAAAACGAGTATATACAGAGGCAAGAGAGAGAGAGAGAGAG 2880
Db 2821 GAAAACTTACAAATTTGAAACAAAACGAGTATATACAGAGGCAAGAGAGAGAGAGAGAGAG 2880
QY 2881 TTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
Db 2881 TTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
QY 2941 GCGGAGATAAACTTGTTCATCGAATTCGAGAGGCTTATCTGTCAGAAATTTATCTGTTATC 3000
Db 2941 GCGGAGATAAACTTGTTCATCGAATTCGAGAGGCTTATCTGTCAGAAATTTATCTGTTATC 3000
QY 3001 CCGGCTGTAATGCGGAAATTTTCAAGAAATTTAGAAAGTTCGATTTATCACTGCAATCTCC 3060
Db 3001 CCGGCTGTAATGCGGAAATTTTCAAGAAATTTAGAAAGTTCGATTTATCACTGCAATCTCC 3060
CTATACGATGCGAGAAATGCTGTTAAAAATGCTGATTTTAAATGATGATGATGCTGCTG 3120

Db 3061 CTATACGATGCGAGAAATGCTGTTAAAAATGCTGATTTTAAATGATGATGATGCTGCTG 3120
QY 3121 AATGTTAAAGGGCATGTAGATGTACACAGAGCCATCACCGCTTCTGCTCTGTTATCCCA 3180
Db 3121 AATGTTAAAGGGCATGTAGATGTACACAGAGCCATCACCGCTTCTGCTCTGTTATCCCA 3180
QY 3181 GAATGGGAAGCAGAAAGTGTCAACAGAGAGTTCGCGTCTGTCCGGGCGTGGCTATATCTC 3240
Db 3181 GAATGGGAAGCAGAAAGTGTCAACAGAGAGTTCGCGTCTGTCCGGGCGTGGCTATATCTC 3240
QY 3241 CGTGTCAAGCGTACAAAGAGGGATATGGAGGGTGTGTAAACGATCCATGAAATCGAG 3300
Db 3241 CGTGTCAAGCGTACAAAGAGGGATATGGAGGGTGTGTAAACGATCCATGAAATCGAG 3300
QY 3301 AACCAATCAGACGAACTTAAAAATTTAAAACTGTGAAGAGAGGAGTGTATCCACCGAT 3360
Db 3301 AACCAATCAGACGAACTTAAAAATTTAAAACTGTGAAGAGAGGAGTGTATCCACCGAT 3360
QY 3361 ACAGAAACGTGTAATGATTTATCTGCACACCAAGGTACAGCATGTATTAATTTCCGTAAT 3420
Db 3361 ACAGAAACGTGTAATGATTTATCTGCACACCAAGGTACAGCATGTATTAATTTCCGTAAT 3420
QY 3421 GCTGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTAAATTAACAAACCGACT 3480
Db 3421 GCTGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTAAATTAACAAACCGACT 3480
QY 3481 TATGAAGAGAGAAACGTATACAGATGTACAGAGATTAATCAATTTGATATGACAGAGGG 3540
Db 3481 TATGAAGAGAGAAACGTATACAGATGTACAGAGATTAATCAATTTGATATGACAGAGGG 3540
QY 3541 TATGTGAATTTATCCACCACTACCAAGTGTGTATATGACAAAGAAATTAAGATTTCCCA 3600
Db 3541 TATGTGAATTTATCCACCACTACCAAGTGTGTATATGACAAAGAAATTAAGATTTCCCA 3600
QY 3601 GAAACCGAATAGGTATGATGAGATTGAGATTGGAGAAACGGAAGGAGATTTATTTAGACAGC 3660
Db 3601 GAAACCGAATAGGTATGATGAGATTGAGATTGGAGAAACGGAAGGAGATTTATTTAGACAGC 3660
QY 3661 GTGGAATTTCTCTTTATGGAGGAATAG 3687
Db 3661 GTGGAATTTCTCTTTATGGAGGAATAG 3687
RESULT 2
AR359382
LOCUS AR359382 4173 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 37 from patent US 6593293.
ACCESSION AR359382
VERSION AR359382.1 GI:33765692
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4173)
AUTHORS Baum,J.A., Chu,C.-R., Donovan,W.P., Gilmer,A.J. and Rupar,M.J.
TITLE Lepidopteran-active Bacillus thuringiensis .delta.-endotoxin
JOURNAL compositions and methods of use
FEATURES Patent: US 6593293-A 37 15-JUL-2003;
Location/Qualifiers
source 1. 4173
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.4%; Score 3666.2; DB 6; Length 4173;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3674; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 TTGACTTCAATAGCAAAATGAGAAATGAGAAATTAATGCTTTATCGATTTCCAGCTGTA 60
Db 1 TTGACTTCAATAGCAAAATGAGAAATGAGAAATTAATGCTTTATCGATTTCCAGCTGTA 60

1141 AATCTGTAGATTATCATTTCTCTCGAGAGTAATTATGAGTGAATCATATGACGGA 1200
1201 GTGCTTCTATGGGAAATTTACCTTGAACCTATTTCATGCTCCCTACTGTTAGATTTAAT 1260
1201 GTGCTTCTATGGGAAATTTACCTTGAACCTATTTCATGCTCCCTACTGTTAGATTTAAT 1260
1261 TTTAGGAACCCCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
1261 TTTAGGAACCCCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
1321 TCACCTGGGCTTCAATTTAAAGATTCAGAAAATGAAATACCCAGAAAACACAGAACGA 1380
1321 TCACCTGGGCTTCAATTTAAAGATTCAGAAAATGAAATACCCAGAAAACACAGAACGA 1380
1381 CCAAAATATGAATCATATATAGTCAATAGGTTATCTCACAATAGGCTCATTTCAAAATCTAGG 1440
1381 CCAAAATATGAATCATATATAGTCAATAGGTTATCTCACAATAGGCTCATTTCAAAATCTAGG 1440
1441 GTGCATGTACAGTATATTTCTTGGACGACCGTAGTGAGATCGTACAAATACCAATTAGT 1500
1441 GTGCATGTACAGTATATTTCTTGGACGACCGTAGTGAGATCGTACAAATACCAATTAGT 1500
1501 TCAGATAGCATAAACAAATACCAATGCTGTAATATCAATCAACCTTAATTCAGGTACCTCT 1560
1501 TCAGATAGCATAAACAAATACCAATGCTGTAATATCAATCAACCTTAATTCAGGTACCTCT 1560
1501 TCAGATAGCATAAACAAATACCAATGCTGTAATATCAATCAACCTTAATTCAGGTACCTCT 1560
1561 GTAGTCAGTGGGCCAGGATTTTACAGGGGGATATATCCGAACTAACGTTAATGAGT 1620
1561 GTAGTCAGTGGGCCAGGATTTTACAGGGGGATATATCCGAACTAACGTTAATGAGT 1620
1561 GTAGTCAGTGGGCCAGGATTTTACAGGGGGATATATCCGAACTAACGTTAATGAGT 1620
1621 GTACTAAGTATGGTCTTAATTTAATTAATATCATCATATACAGGGATCGGTGAGAGTT 1680
1621 GTACTAAGTATGGTCTTAATTTAATTAATATCATCATATACAGGGATCGGTGAGAGTT 1680
1621 GTACTAAGTATGGTCTTAATTTAATTAATATCATCATATACAGGGATCGGTGAGAGTT 1680
1681 CGTTATGCTCTCTCAACCAATGCTGAGGGTAACTGTCGGAGGAGTACTACTTTT 1740
1681 CGTTATGCTCTCTCAACCAATGCTGAGGGTAACTGTCGGAGGAGTACTACTTTT 1740
1681 CGTTATGCTCTCTCAACCAATGCTGAGGGTAACTGTCGGAGGAGTACTACTTTT 1740
1741 GATCAAGATTTCCCTAGTACTATGATGCAATGATCTTTGACATCTCAATCATTTAGA 1800
1741 GATCAAGATTTCCCTAGTACTATGATGCAATGATCTTTGACATCTCAATCATTTAGA 1800
1741 GATCAAGATTTCCCTAGTACTATGATGCAATGATCTTTGACATCTCAATCATTTAGA 1800
1801 TTTGAGAAATTTCTGTAGGTATTAGTGCATCTGGGAGTCAAACTGCTGGAAATAGTATA 1860
1801 TTTGAGAAATTTCTGTAGGTATTAGTGCATCTGGGAGTCAAACTGCTGGAAATAGTATA 1860
1801 TTTGAGAAATTTCTGTAGGTATTAGTGCATCTGGGAGTCAAACTGCTGGAAATAGTATA 1860
1861 AGTAATATGCAAGTATGAGTAACTGATGCAATGATCTTTGACATCTCAATCATTTAGA 1920
1861 AGTAATATGCAAGTATGAGTAACTGATGCAATGATCTTTGACATCTCAATCATTTAGA 1920
1861 AGTAATATGCAAGTATGAGTAACTGATGCAATGATCTTTGACATCTCAATCATTTAGA 1920
1921 GCAACCTTCGAGGAGAAATACGATTTAGAAAGGGGCGAGAGGGGCGGTGATGCTCTGTTT 1980
1921 GCAACCTTCGAGGAGAAATACGATTTAGAAAGGGGCGAGAGGGGCGGTGATGCTCTGTTT 1980
1921 GCAACCTTCGAGGAGAAATACGATTTAGAAAGGGGCGAGAGGGGCGGTGATGCTCTGTTT 1980
1981 ACTAATACGAATCAAGAGATTTGAAACAGATGTGACAGATTTATCATTTGATCAAGTA 2040
1981 ACTAATACGAATCAAGAGATTTGAAACAGATGTGACAGATTTATCATTTGATCAAGTA 2040
1981 ACTAATACGAATCAAGAGATTTGAAACAGATGTGACAGATTTATCATTTGATCAAGTA 2040
2041 TCCAAATTTAGTGGGCTGTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2100
2041 TCCAAATTTAGTGGGCTGTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2100
2041 TCCAAATTTAGTGGGCTGTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2100
2101 GAGAAAGTGAATATATGCGAAACGATCACTGATGAGAAAGAACTTACTCCAGATCCAAAC 2160
2101 GAGAAAGTGAATATATGCGAAACGATCACTGATGAGAAAGAACTTACTCCAGATCCAAAC 2160
2101 GAGAAAGTGAATATATGCGAAACGATCACTGATGAGAAAGAACTTACTCCAGATCCAAAC 2160
2161 TCCAAATTTAGTGGGCTGTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2220
2161 TCCAAATTTAGTGGGCTGTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2220
2161 TCCAAATTTAGTGGGCTGTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2220
2221 TCTATCCATGAACTCAATCTCAATGATGAGTGGGAGTGAACATTTACAAATCCAGGAA 2280

61 TCGAATCATTTCCACAAATGATCTATACAGATGCTCGTATTGAGATTTCTTTGTT 120
61 TCGAATCATTTCCACAAATGATCTATACAGATGCTCGTATTGAGATTTCTTTGTT 120
121 ATAGCCGAGGGAATTAATCAATCCATCTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
121 ATAGCCGAGGGAATTAATCAATCCATCTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
121 AACATAGCTGTTAGATTAATCAATCCATCTGTTAGCGCATCAACAGTCCAAACGGGTATT 240
121 AACATAGCTGTTAGATTAATCAATCCATCTGTTAGCGCATCAACAGTCCAAACGGGTATT 240
181 AACATAGCTGTTAGATTAATCAATCCATCTGTTAGCGCATCAACAGTCCAAACGGGTATT 300
241 TTTTATAGTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 360
241 TTTTATAGTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 360
301 CTAGAACATGTCGAACCAATTAATCAACAAATTAACAGAAATGCTTAGGAATACGCA 360
301 CTAGAACATGTCGAACCAATTAATCAACAAATTAACAGAAATGCTTAGGAATACGCA 360
361 CTGCTCGAATTAACAAAGTTTATAGGATCTCTTTAGAGCTTATCAACAGTCACTTGAAGAT 420
361 CTGCTCGAATTAACAAAGTTTATAGGATCTCTTTAGAGCTTATCAACAGTCACTTGAAGAT 420
421 TGCTAGAAACCGGTGATGTCGAAGAACGAGAGTGTCTTTATACCCATATATAGCC 480
421 TGCTAGAAACCGGTGATGTCGAAGAACGAGAGTGTCTTTATACCCATATATAGCC 480
481 TTAGAACCTGATTTCTTAATCGGATGCGCTTTTTCGCAATTTAGAAACCAAGAGTTCCA 540
481 TTAGAACCTGATTTCTTAATCGGATGCGCTTTTTCGCAATTTAGAAACCAAGAGTTCCA 540
481 TTAGAACCTGATTTCTTAATCGGATGCGCTTTTTCGCAATTTAGAAACCAAGAGTTCCA 540
541 TTATTAATGATGATGCTCAAGTGCAGAAATTTACACCTATTAATTTAGAGATGCTCTCT 600
541 TTATTAATGATGATGCTCAAGTGCAGAAATTTACACCTATTAATTTAGAGATGCTCTCT 600
541 TTATTAATGATGATGCTCAAGTGCAGAAATTTACACCTATTAATTTAGAGATGCTCTCT 600
601 CTTTTTGTAGTGAATTTGGGCTTACATCGAGAAATTTCAACGTTATTTAGGGGCCAA 660
601 CTTTTTGTAGTGAATTTGGGCTTACATCGAGAAATTTCAACGTTATTTAGGGGCCAA 660
601 CTTTTTGTAGTGAATTTGGGCTTACATCGAGAAATTTCAACGTTATTTAGGGGCCAA 660
661 GTGGAAACAAACGAGATATTTCGCAATTTGGTGAATGATTAATACAGTCTTAAT 720
661 GTGGAAACAAACGAGATATTTCGCAATTTGGTGAATGATTAATACAGTCTTAAT 720
661 GTGGAAACAAACGAGATATTTCGCAATTTGGTGAATGATTAATACAGTCTTAAT 720
721 AGCTTGAAGGGAACAAATGCGCAAGTTGGGTGCTTATATCAATCTCGGATGATCTA 780
721 AGCTTGAAGGGAACAAATGCGCAAGTTGGGTGCTTATATCAATCTCGGATGATCTA 780
721 AGCTTGAAGGGAACAAATGCGCAAGTTGGGTGCTTATATCAATCTCGGATGATCTA 780
781 ACCTTGAAGGGAACAAATGCGCAAGTTGGGTGCTTATATCAATCTCGGATGATCTA 840
781 ACCTTGAAGGGAACAAATGCGCAAGTTGGGTGCTTATATCAATCTCGGATGATCTA 840
781 ACCTTGAAGGGAACAAATGCGCAAGTTGGGTGCTTATATCAATCTCGGATGATCTA 840
841 ATAAATACGAGTGTCTCAAGTTTAAACAGGAAATTTATACAGCGCAATTTGAGCAACAGG 900
841 ATAAATACGAGTGTCTCAAGTTTAAACAGGAAATTTATACAGCGCAATTTGAGCAACAGG 900
841 ATAAATACGAGTGTCTCAAGTTTAAACAGGAAATTTATACAGCGCAATTTGAGCAACAGG 900
901 GTAAATATGCAAGTATGAATTTGATTAATTAATGCACTTTGCTTTTCCGCTATAGAG 960
901 GTAAATATGCAAGTATGAATTTGATTAATTAATGCACTTTGCTTTTCCGCTATAGAG 960
901 GTAAATATGCAAGTATGAATTTGATTAATTAATGCACTTTGCTTTTCCGCTATAGAG 960
961 ACTGCGGTTATCCGAAGCCGCTACTAGGATATGATTTTCTAGCAACACTTACAAATTTTAC 1020
961 ACTGCGGTTATCCGAAGCCGCTACTAGGATATGATTTTCTAGCAACACTTACAAATTTTAC 1020
961 ACTGCGGTTATCCGAAGCCGCTACTAGGATATGATTTTCTAGCAACACTTACAAATTTTAC 1020
1021 ACTTTCATCAGATGAGTGTCTACTAGGATATGATTTTCTAGCAACACTTACAAATTTTAC 1080
1021 ACTTTCATCAGATGAGTGTCTACTAGGATATGATTTTCTAGCAACACTTACAAATTTTAC 1080
1021 ACTTTCATCAGATGAGTGTCTACTAGGATATGATTTTCTAGCAACACTTACAAATTTTAC 1080
1081 TCTCGGCAATAGGAGCGGATTAATATCTCAACGCTGCTTACCAATTTCTATT 1140
1081 TCTCGGCAATAGGAGCGGATTAATATCTCAACGCTGCTTACCAATTTCTATT 1140
1081 TCTCGGCAATAGGAGCGGATTAATATCTCAACGCTGCTTACCAATTTCTATT 1140
1141 AATCTGTAGATTATCATTTCTCTCGAGAGTATATGAGTGAATCATATGACGGA 1200

2221	Db	TCTATCCATGAA	CAATCTGAACTATGGATATGGTGGGGAAGTGGAGAACATTA	CAATCCAGGAA	2288
2281	Qy	GGAAATGACGTATTTAAAGAGAA	TTACGTACACTACCGGGGACCTTTAAATAGAGTGTAT	2340	
2281	Db	GGAAATGACGTATTTAAAGAGAA	TTACGTACACTACCGGGGACCTTTAAATAGAGTGTAT	2340	
2341	Qy	CCGACGTATTTATCAAAAATAGGAGAGT	CGGAATTAAGAGTTATATCTCGTACCAA	2400	
2341	Db	CCGACGTATTTATCAAAAATAGGAGAGT	CGGAATTAAGAGTTATATCTCGTACCAA	2400	
2401	Qy	TTAAGAGGCTATATTTGAAGATAGT	CAAGATTTAGAGATATATTTGATTCGTATAATGCG	2460	
2401	Db	TTAAGAGGCTATATTTGAAGATAGT	CAAGATTTAGAGATATATTTGATTCGTATAATGCG	2460	
2461	Qy	AAAATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGTATGGCGCTTTTCAGTTGAAAGC	2520		
2461	Db	AAAATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGTATGGCGCTTTTCAGTTGAAAGC	2520		
2521	Qy	CCAAATCGGAAGTTCGGAGAACCGAATCGAATCGATCGGCACCA	CATTTTGAATGGAATCCGTGAT	2580	
2521	Db	CCAAATCGGAAGTTCGGAGAACCGAATCGAATCGAATCGGCACCA	CATTTTGAATGGAATCCGTGAT	2580	
2581	Qy	CTAGATGTTCTTCGACAGATGGAGAAAATGTGCGCATCATTTCCCATCATTTCTCTTTG	2640		
2581	Db	CTAGATGTTCTTCGACAGATGGAGAAAATGTGCGCATCATTTCCCATCATTTCTCTTTG	2640		
2641	Qy	GATATTGATATTTGGATGACACAGACTTCGATGAGAAATCTAGCGGTGCGGTGGTATTCAG	2700		
2641	Db	GATATTGATATTTGGATGACACAGACTTCGATGAGAAATCTAGCGGTGCGGTGGTATTCAG	2700		
2701	Qy	ATTAAGACCGGAAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAACCA	2760		
2701	Db	ATTAAGACCGGAAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAACCA	2760		
2761	Qy	TTATTAGGAGAGCACTGTCTGTGTGCAAGAGCAGAGAGAAAATA	TGGAGAGACAAACGT	2820	
2761	Db	TTATTAGGAGAGCACTGTCTGTGTGCAAGAGCAGAGAGAAAATA	TGGAGAGACAAACGT	2820	
2821	Qy	GAAAACTCAATTTGGAAA	CAAAACGAGTATATACAGAGCAGAAAAGAGCTGTGGATGCT	2880	
2821	Db	GAAAACTCAATTTGGAAA	CAAAACGAGTATATACAGAGCAGAAAAGAGCTGTGGATGCT	2880	
2881	Qy	TTATTTGTAGATTTCAATATAATAGATTACAAGCGGATACAACATTTGCATGATTCAT	2940		
2881	Db	TTATTTGTAGATTTCAATATAATAGATTACAAGCGGATACAACATTTGCATGATTCAT	2940		
2941	Qy	CGCGCAGATAAATCTGTTTCATCGAATTTGAGAGGCTTATCTGTGAGAAATATCTGTTATC	3000		
2941	Db	CGCGCAGATAAATCTGTTTCATCGAATTTGAGAGGCTTATCTGTGAGAAATATCTGTTATC	3000		
3001	Qy	CCGGTGTAAATGCGGAAAATTTTTTGAAGAAATTAGAAGTTCGATATACCTGCAATCTCC	3060		
3001	Db	CCGGTGTAAATGCGGAAAATTTTTTGAAGAAATTAGAAGTTCGATATACCTGCAATCTCC	3060		
3061	Qy	CTATACGATCGGAAAATGTCGTTTAAAAATGGTGATTTAAATATGGAATTAGCATGCTGG	3120		
3061	Db	CTATACGATCGGAAAATGTCGTTTAAAAATGGTGATTTAAATATGGAATTAGCATGCTGG	3120		
3121	Qy	AATGTAAGGGCATGTAGATGTATCAACAGAGCCATCACCGTGTCTGCTTGTATATCCCA	3180		
3121	Db	AATGTAAGGGCATGTAGATGTATCAACAGAGCCATCACCGTGTCTGCTTGTATATCCCA	3180		
3181	Qy	GAATGGGAAGCAGAAAGTGTACAGCAGTTCGCGTCTGTCGCGGGCGTGCTATATCCCTC	3240		
3181	Db	GAATGGGAAGCAGAAAGTGTACAGCAGTTCGCGTCTGTCGCGGGCGTGCTATATCCCTC	3240		
3241	Qy	CGTGTCAACGCTACAAAGGGGATATGGAGAGGGTTGTGTAAACGATCCATGAATCGAG	3300		
3241	Db	CGTGTCAACGCTACAAAGGGGATATGGAGAGGGTTGTGTAAACGATCCATGAATCGAG	3300		
3301	Qy	AACAAATCAGACGAACTAAATTTAAAACTGTGTAAGAGAGGAGTGTATCCAAACGGAT	3360		
3301	Db	AACAAATCAGACGAACTAAATTTAAAACTGTGTAAGAGAGGAGTGTATCCAAACGGAT	3360		

Qy	3361	ACAGGAACGTTGTAAATGATTATATCGCACCAAGGTACAGCAGTATGTAAATTCCTCCGTAAT	3420
Dd	3361	ACAGGAACGTTGTAAATGATTATATCGCACCAAGGTACAGCAGTATGTAAATTCCTCCGTAAT	3420
Qy	3421	GCTGGATATCAGGATGCATATGAAGTTTGATCTACTACAGCATCTCTTAAATTCACAACCGACT	3480
Dd	3421	GCTGGATATCAGGATGCATATGAAGTTTGATCTACTACAGCATCTCTTAAATTCACAACCGACT	3480
Qy	3481	TATGAAGAAGAACGTTATACAGATGTACGAAGAGATAAATCATTTGTGAATATGACAGAGG	3540
Dd	3481	TATGAAGAAGAACGTTATACAGATGTACGAAGAGATAAATCATTTGTGAATATGACAGAGG	3540
Qy	3541	TATGTGAATATTCACCACTACAGCTGTGTTATATGACAAAAGAAATTAGAATATCTCCCA	3600
Dd	3541	TATGTGAATATTCACCACTACAGCTGTGTTATATGACAAAAGAAATTAGAATATCTCCCA	3600
Qy	3601	GAACCCGATGAAGTATGATTTGAGATTTGGAGAACCGNAGGGAAGTTATTGTAGACAGC	3660
Dd	3601	GAACCCGATGAAGTATGATTTGAGATTTGGAGAACCGNAGGGAAGTTATTGTAGACAGC	3660
Qy	3661	GTGGAATTACTCTCTTATGGAGGAATAG	3687
Dd	3661	ATAGAATTACTCTCTTATGGAGGAATAG	3687

RESULT 3
AX098667

LOCUS AX098667 4173 bp DNA linear PAT 02-APR-2001

DEFINITION Sequence 37 from Patent WO0119859.

ACCESSION AX098667

VERSION AX098667.1 GI:13537912

KEYWORDS

SOURCE Bacillus thuringiensis

ORGANISM Bacillus thuringiensis

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE 1

AUTHORS Baum, J.A., Chu, C.R., Donovan, W.P., Gilmer, A.J. and Rupar, M.J.

TITLE Lepidopteran-active bacillus thuringiensis delta-endotoxin compositions and methods of use

JOURNAL Patent: WO 0119859-A 37 22-MAR-2001;

MONSANTO COMPANY (US)

FEATURES source Location/Qualifiers

source 1..4173

/organism="Bacillus thuringiensis"

/mol_type="unassigned DNA"

/db_xref="taxon:1428"

1..3687

/notes="unnamed protein product"

CDS

ORIGIN		99.4%; Score 3666.2; DB 6; Length 4173;	
Query Match		99.6%; Pred. No. 0;	
Beet Local Similarity		0; Mismatches 13; Indels 0; Gaps 0;	
Matches 3674; Conservative			
		YGEVGVTHIEENTNDELKFNCEBEVYPTDTGTCNDYTAHQSTAAACNSRNAGYEDA	
		YEVDTTASVNYKPYBESETYTDVRDNHCEYDRGVVNYPPLPAGYVTKELSYFPETDK	
		VWIEIGETEGTFIVDSIELLIMEE"	
QY	1	TTGACTTCAAAATAGGAAAAATGAGAAATGAATTAATAATTAATAATGCTTTATCGATTCAGCTGTA	60
DB	1	TTGACTTCAAAATAGGAAAAATGAGAAATGAATTAATAATGCTTTATCGATTCAGCTGTA	60
QY	61	TCGAATCAATCCACAAATGGAATCTATCAACAGATGCTCGTATGAGGATCTTTGTGT	120
DB	61	TCGAATCAATCCACAAATGGAATCTATCAACAGATGCTCGTATGAGGATCTTTGTGT	120
QY	121	ATAGCGGAGGGAATAATATCAATCCACTTTGAGGCATCAACAGTCCAAACGGTATT	180
DB	121	ATAGCGGAGGGAATAATATCAATCCACTTTGAGGCATCAACAGTCCAAACGGTATT	180
QY	181	AACTAGCTGGTAGAATACTAGGTGTATTAGGCGTACCGTTTCTGCGACAAATAGCTAGT	240
DB	181	AACTAGCTGGTAGAATACTAGGTGTATTAGGCGTACCGTTTCTGCGACAAATAGCTAGT	240
QY	241	TTTTATAGTTTCTTGTTGGTGAATATGCGCCCGCGCAGAGATCAGTGGGAAATTTTC	300
DB	241	TTTTATAGTTTCTTGTTGGTGAATATGCGCCCGCGCAGAGATCAGTGGGAAATTTTC	300
QY	301	CTAGNACATGTCGAAACAACTTATAATCAACAAATACAGAAATGCTAGGAATACGGCA	360
DB	301	CTAGNACATGTCGAAACAACTTATAATCAACAAATACAGAAATGCTAGGAATACGGCA	360
QY	361	CTTGCTCGATTACAGGTTTGGAGATTCCTTTAGAGCTTATCAACAGTCACTTGAAGAT	420
DB	361	CTTGCTCGATTACAGGTTTGGAGATTCCTTTAGAGCTTATCAACAGTCACTTGAAGAT	420
QY	421	TGGCTAGAAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATACCCAAATATATAGCC	480
DB	421	TGGCTAGAAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATACCCAAATATATAGCC	480
QY	481	TTAGAACTTGATTTCTTAAATGCGATGCGCTTTTCGCAATTAAGAACCCAGAGTTCCA	540
DB	481	TTAGAACTTGATTTCTTAAATGCGATGCGCTTTTCGCAATTAAGAACCCAGAGTTCCA	540
QY	541	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTTATTAATTTGAGAGTGCCTCT	600
DB	541	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTTATTAATTTGAGAGTGCCTCT	600
QY	601	CTTTTTCGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACCGTATTATGAGCGCCAA	660
DB	601	CTTTTTCGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACCGTATTATGAGCGCCAA	660
QY	661	GTGGAAACAAACGAGAGATTTCCGACTATTGCTAGAAATGGTATATAACAGGCTTAAAT	720
DB	661	GTGGAAACAAACGAGAGATTTATCCGACTATTGCTAGAAATGGTATATAACAGGCTTAAAT	720
QY	721	AGCTTTGAGAGGGAACAAATGCGGAAAGTTGGTGGTGTATTAATCAATTCGGTAGAGATCTA	780
DB	721	AGCTTTGAGAGGGAACAAATGCGGAAAGTTGGTGGTGTATTAATCAATTCGGTAGAGATCTA	780
QY	781	ACGTTAGGAGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA	840
DB	781	ACGTTAGGAGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA	840
QY	841	ATAAATACGAGTGTCAAGTAAACAGGGAAGTTTATACAGCGCAATTTGAGCAACAGG	900
DB	841	ATAAATACGAGTGTCAAGTAAACAGGGAAGTTTATACAGCGCAATTTGAGCAACAGG	900
QY	901	GTAATATGCAAGTATGAATTTGGTATAATAATATGCACTTCTGTTTCCGCTATAGAG	960
DB	901	GTAATATGCAAGTATGAATTTGGTATAATAATATGCACTTCTGTTTCCGCTATAGAG	960

QY	961	ACTGGGTTATCCGAAGCCCGCATCTACTTGATTTTCTAGAACAACTTACAAATTTTATAGC	1020
DB	961	ACTGGGTTATCCGAAGCCCGCATCTACTTGATTTTCTAGAACAACTTACAAATTTTATAGC	1020
QY	1021	ACTTCATCAGATGAGTGTCTAGGCATATGACTTACTTGGGGGGGACACAATTTCAA	1080
DB	1021	ACTTCATCAGATGAGTGTCTAGGCATATGACTTACTTGGGGGGGACACAATTTCAA	1080
QY	1081	TCCTGSCCAATAGGAGGCGGATTAATACCTACGCAATGGTCTACCAATATCTTCTATT	1140
DB	1081	TCCTGSCCAATAGGAGGCGGATTAATACCTACGCAATGGTCTACCAATATCTTCTATT	1140
QY	1141	AATCTGTTAGATTTATCTTCTCTCGAGACGTFATTTGGACTGAATCATATGACGGA	1200
DB	1141	AATCTGTTAGATTTATCTTCTCTCGAGACGTFATTTGGACTGAATCATATGACGGA	1200
QY	1201	GTGCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGCCCTACTGTTAGATTTAAT	1260
DB	1201	GTGCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGCCCTACTGTTAGATTTAAT	1260
QY	1261	TTTAGGAACCTCAGAAATCTTTTGAAGAGGTFATGCTTAACATATAGTCAACCTTATGAG	1320
DB	1261	TTTAGGAACCTCAGAAATCTTTTGAAGAGGTFATGCTTAACATATAGTCAACCTTATGAG	1320
QY	1321	TCACCTGGGCTTCAATTTAAAGATTTCAAGAACTGAAATTTACCAAGAAACACAGACGA	1380
DB	1321	TCACCTGGGCTTCAATTTAAAGATTTCAAGAACTGAAATTTACCAAGAAACACAGACGA	1380
QY	1381	CGAAATATGAATCATATAGTATAGTGTATCTACATAGGGCTCATTTCACAATCTAGG	1440
DB	1381	CGAAATATGAATCATATAGTATAGTGTATCTACATAGGGCTCATTTCACAATCTAGG	1440
QY	1441	GTGCATGTACAGTATATTTTGGACGCAACGCTAGTCAGATCGTACAAATACCATTTAGT	1500
DB	1441	GTGCATGTACAGTATATTTTGGACGCAACGCTAGTCAGATCGTACAAATACCATTTAGT	1500
QY	1501	TCAGATAGCATAACACAAATACCAATTTGGTAAATCATTTCAACCTTTAAATTCAGGTACCTCT	1560
DB	1501	TCAGATAGCATAACACAAATACCAATTTGGTAAATCATTTCAACCTTTAAATTCAGGTACCTCT	1560
QY	1561	GTAGTCAAGTGGCCAGGATTTTACAGGAGGGAATTAATCCGAACTAACTGTTAATGGTAGT	1620
DB	1561	GTAGTCAAGTGGCCAGGATTTTACAGGAGGGAATTAATCCGAACTAACTGTTAATGGTAGT	1620
QY	1621	GTACTAAGTATGGGTCTTAAATTAATATACATCATTTACAGCGGTATCGGTGAGGTT	1680
DB	1621	GTACTAAGTATGGGTCTTAAATTAATATACATCATTTACAGCGGTATCGGTGAGGTT	1680
QY	1681	CGTTATGCTGCTTCTCAAAACAAATGGTCTGAGGGTAACTGTCGAGGGAGTACTACTTTT	1740
DB	1681	CGTTATGCTGCTTCTCAAAACAAATGGTCTGAGGGTAACTGTCGAGGGAGTACTACTTTT	1740
QY	1741	GATCAAGGATTCCTTAGTACTATGATGCAAAATGATCTTTGACATCTCAATCAATTTAGA	1800
DB	1741	GATCAAGGATTCCTTAGTACTATGATGCAAAATGATCTTTGACATCTCAATCAATTTAGA	1800
QY	1801	TTTCGAGAAATTTCTGTAGGTATTTAGTCACTCTGCGAGTCAAACTGCTGGAATAGTATA	1860
DB	1801	TTTCGAGAAATTTCTGTAGGTATTTAGTCACTCTGCGAGTCAAACTGCTGGAATAGTATA	1860
QY	1861	AGTAATAATGAGTACGAGAAAGGTTTCACTTTGATAAAATTTGAATTTCAATTTCAATTTACT	1920
DB	1861	AGTAATAATGAGTACGAGAAAGGTTTCACTTTGATAAAATTTGAATTTCAATTTCAATTTACT	1920
QY	1921	GCAACTTTCGAAAGCAAGATAGCATTTTGAAGGCGCAAGAGCGGTGAATGCTCTGTTT	1980
DB	1921	GCAACTTTCGAAAGCAAGATAGCATTTTGAAGGCGCAAGAGCGGTGAATGCTCTGTTT	1980
QY	1981	ACTAATATCCGAAATTCGAAAGATTTGAAACAGATGTGACAGATTTATCATTTGATCAAGTA	2040
DB	1981	ACTAATATCCGAAATTCGAAAGATTTGAAACAGATGTGACAGATTTATCATTTGATCAAGTA	2040
QY	2041	TCCAAATTTAGTGGGCTTTTATCGGATGAATTTCTGCTTTAGATGAAGAGAGATTTACTT	2100
DB	2041	TCCAAATTTAGTGGGCTTTTATCGGATGAATTTCTGCTTTAGATGAAGAGAGATTTACTT	2100

Db 2041 TCCAAATTAGTGGCGTGTATTCGATGAATTCCTCTAGATGAAGAGAGAATTACTT 2100
Qy 2101 GAGAAAGTGAATATATGCGAAACGCTCAGTGTATGAAGAAATCTTACTCCAGATCCAAAC 2160
Db 2101 GAGAAAGTGAATATATGCGAAACGCTCAGTGTATGAAGAAATCTTACTCCAGATCCAAAC 2160
Qy 2161 TTCACATCCATCAATTAAGCAACCAAGCTTCTATATCTACTATAGCAATCGAATTTTACA 2220
Db 2161 TTCACATCCATCAATTAAGCAACCAAGCTTCTATATCTACTATAGCAATCGAATTTTACA 2220
Qy 2221 TCTATCCATGAACATCTGAACATGGATGGTGGGAAGTGAGAACATTTACAAATCCAGAA 2280
Db 2221 TCTATCCATGAACATCTGAACATGGATGGTGGGAAGTGAGAACATTTACAAATCCAGAA 2280
Qy 2281 GGAATACGCTATTTAAAGAGAAATPACGTCACTACCGGGGACTTTTAAATGAGTGTAT 2340
Db 2281 GGAATACGCTATTTAAAGAGAAATPACGTCACTACCGGGGACTTTTAAATGAGTGTAT 2340
Qy 2341 CCGACGCTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTTATCTCGCTACCAA 2400
Db 2341 CCGACGCTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTTATCTCGCTACCAA 2400
Qy 2401 TTAAGAGGCTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAATGCG 2460
Db 2401 TTAAGAGGCTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAATGCG 2460
Qy 2461 AAAACATGAACATTTGGATGTTCCAGTACCGAGTCCGATATGCGCGCTTTTCAAGTGAAGC 2520
Db 2461 AAAACATGAACATTTGGATGTTCCAGTACCGAGTCCGATATGCGCGCTTTTCAAGTGAAGC 2520
Qy 2521 CCAATCGGAAGTGGGAGAACCGAATCGATCGCGCACCAATTTTGAATGGAATCCTGAT 2580
Db 2521 CCAATCGGAAGTGGGAGAACCGAATCGATCGCGCACCAATTTTGAATGGAATCCTGAT 2580
Qy 2581 CTAGATTTGCTCGCAGAGATGAGAAATATGTCGCATCATTTCCATCATTTCTCTTTG 2640
Db 2581 CTAGATTTGCTCGCAGAGATGAGAAATATGTCGCATCATTTCCATCATTTCTCTTTG 2640
Qy 2641 GATATTGATATTGGATGCACAGACTTGCATCAGATCTAGCGGTGGGTGATTTCAAG 2700
Db 2641 GATATTGATATTGGATGCACAGACTTGCATCAGATCTAGCGGTGGGTGATTTCAAG 2700
Qy 2701 ATTAAGACGCGAGAGGTCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAACCA 2760
Db 2701 ATTAAGACGCGAGAGGTCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAACCA 2760
Qy 2761 TTATTAGAGAGCACTGCTCTCGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Db 2761 TTATTAGAGAGAGCACTGCTCTCGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Qy 2821 GAAAACTTACAAATGGAACAAACGAGTATATACAGAGGCAAGAGAGAGAGAGAGAGAGAG 2880
Db 2821 GAAAACTTACAAATGGAACAAACGAGTATATACAGAGGCAAGAGAGAGAGAGAGAGAGAG 2880
Qy 2881 TTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
Db 2881 TTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
Qy 2941 GCGGAGAGATAAATTTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTTATC 3000
Db 2941 GCGGAGAGATAAATTTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTTATC 3000
Qy 3001 CCGGGTGAATGCGGAATTTTGAAGAAATTTAGAGGTCGATTTATCAATTCCTCC 3060
Db 3001 CCGGGTGAATGCGGAATTTTGAAGAAATTTAGAGGTCGATTTATCAATTCCTCC 3060
Qy 3061 CTATACGATGCGAGAAATGCTTTAAATGCTGATTTTAAATGAGTATAGCATGCTGG 3120
Db 3061 CTATACGATGCGAGAAATGCTTTAAATGCTGATTTTAAATGAGTATAGCATGCTGG 3120
3121 AATGTAAAGGCGATGTAGATGTACACAGAGCCATCACCGTCTGTCTTGTATATCCCA 3180

Db 3121 AATGTAAAGGCGCATGTAGATGTACACAGAGCCATCACCGTCTGTCTTGTATATCCCA 3180
Qy 3181 GAATGGGAGCAGAAAGTGTCAACAGAGATTTCGCGTCTGTCCGGGCGTGGCTATATCTTC 3240
Db 3181 GAATGGGAGCAGAAAGTGTCAACAGAGATTTCGCGTCTGTCCGGGCGTGGCTATATCTTC 3240
Qy 3241 CGTGTCAAGCGGTACAAAGAGGGATATGGAGAGGTTGTGTAAACGATCCATGAATTCGAG 3300
Db 3241 CGTGTCAAGCGGTACAAAGAGGGATATGGAGAGGTTGTGTAAACGATCCATGAATTCGAG 3300
Qy 3301 AACAAATCAGACGAACTAAATTTAAAACTGTGAAGAGAGGAGTGTATCCAAACGAT 3360
Db 3301 AACAAATCAGACGAACTAAATTTAAAACTGTGAAGAGAGGAGTGTATCCAAACGAT 3360
Qy 3361 ACAGAAACGTTAATGATTTATCTACTGCACCAAGGTACAGCAGTATGTAATTCGCTAAT 3420
Db 3361 ACAGAAACGTTAATGATTTATCTACTGCACCAAGGTACAGCAGTATGTAATTCGCTAAT 3420
Qy 3421 GCTGGATATGAGGATGATATGAAAGTTGATACTACAGCATCTGTTAATTACAAACCGACT 3480
Db 3421 GCTGGATATGAGGATGATATGAAAGTTGATACTACAGCATCTGTTAATTACAAACCGACT 3480
Qy 3481 TATGAAGAGAAACGTTATACAGATGTACGAAGAGATTAATCAATTTGAATATGACAGAGG 3540
Db 3481 TATGAAGAGAAACGTTATACAGATGTACGAAGAGATTAATCAATTTGAATATGACAGAGG 3540
Qy 3541 TATGTGAATTTATCCACCATACAGCTGTTTATATGACAAAGAAATAGATATCTTCCCA 3600
Db 3541 TATGTGAATTTATCCACCATACAGCTGTTTATATGACAAAGAAATAGATATCTTCCCA 3600
Qy 3601 GAAACCGAATAGGTATGAGATTGAGATTGGAGAAACGGAAGGAAATTTATTGTAGACAGC 3660
Db 3601 GAAACCGAATAGGTATGAGATTGAGATTGGAGAAACGGAAGGAAATTTATTGTAGACAGC 3660
Qy 3661 GTGGAATTTACTCTTTATGGAGGAATAG 3687
Db 3661 ATAGAATTTACTCTTTATGGAGGAATAG 3687

RESULT 4
AF077326 3684 bp DNA linear BCT 26-AUG-1998
DEFINITION Bacillus thuringiensis CryI_{BeI} delta-endotoxin gene, complete cds.
ACCESSION AF077326
VERSION AF077326.1 GI:3360518
KEYWORDS
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 3684)
Payne, J., Cummings, D. A., Cannon, R. J. C., Narva, K. E. and Stelman, S. J.
Bacillus thuringiensis genes encoding lepidopteran-active toxins
Patent: US 5723758 03-MAR-1998;
Mycogen Corporation; San Diego, CA
REFERENCE 2 (bases 1 to 3684)
Stelman, S. J.
Direct Submission
Submitted (08-JUL-1998) Molecular Biology, Mycogen Corporation,
5501 Oberlin Dr., San Diego, CA 92121, USA
Location/Qualifiers
1. 3684
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="PS158C2"
/db_xref="taxon:1428"
/plasmid="pMYC2383"
1. 3684
/codon_start=1
/transl_table=11
/product="CryI_{BeI} delta-endotoxin"
/protein_id="AAC32850.1"
/db_xref="GI:3360519"

FEATURES
source
CDS

/translation="MTRNKNENEIINALSIPAVSNHSAQNMILSTDARIEDSLCIAEG
 NNIDPVSASTVQTGINIAGRIILGVLPFAGQIASFYSLVGLMPPRGDRPWEIFLE
 HVQLIRQQVTEINTRDTALARLQGLGNSFRAYQOQSLIEDLWENRDLARTSLVLYQYIA
 LEDFLNMLPFAIRNOEPLLMVYQAANLHLLLRDLSPGSEFLTQSEIORYYE
 ROVKETREYSDCARWYNTDGLNRGTNAESMLRWYQFRDLTLGLVDLVALPPSYDT
 RVYPMNTSAQLTREIYTDPIGRNAPSGPASTWPNFNAPSPSAIEAAVIRPPLHLD
 PEOLTIFSVLSRNSNTQYNNWYGHLESTIRGSLSTHGNNTNTSINPVLQPTSR
 DVYRTSEFAGINILLTTPVNGVWPAPRNWRNPLNSLRGSLLYTIGYTGVTQLFDBET
 ELPEPTERNPYESYSHRUSNIRLISGNTLRAPVSWTHRSADRTNTISSDLSITPL
 VKSFLNLSGTSVYGGFTGGDIIRTNVSGSVLSMGLNFNTSLQRYRVRYAASQT
 MVLRTVGGSTTDDQGPSPMSANESLTSQSFPAEPFVIGSAGSQTAGISISNAG
 RQTFPHDKBFIPIATFAEYDLERAQAVNALFTNPNRRLKTVGVYDHIDEVSNL
 VACLDEFCLDERLELLEKVKYAKRLSDERNLQDPNTSINKQDPFI STNQSNTFS
 IHQSHGHWGNSNIIQEGNDVPKENYVILPGTFNECYPTLYLQKIGBAELKAYTRY
 QLSGYIEDSODLSIYLIRNAKHTLDPGTSVWPLSVESPIGRCEBPNKAPHFEW
 NPDLDCSDEKXCAHSHHFSLDIDVGCIDLHENLGVWVVFKIOEGBHARLGNLF
 IEBKPLLGALSARAEKRWDRKQLQLETKRVYTEAKEADVALFVDSQYDRLOAD
 TNIGMTHAADKLVHRIEAYLSBELSPGVNABRI FEELEGRIITAI SLVDARNVKNQ
 DPNNGLACWNVKGVHDVQOQSHRPSVLVIPEWEAEVSQAVRVCPRGVILRVYAKEGY
 GESCVTIHEIENNTDELKPKNCBEEBIVYPTDGTGTCNDYTAHQGTACNSRNAGYEDAY
 EVDITASVNTKPTYBEEYTDVRRDNHCEYDRGVYVPPVPAGYMTKELEYIFPETDKV
 WIEBIGTEGKFI VDSVELLMEE"

ORIGIN

Query Match 88.6%; Score 3266.6; DB 1; Length 3684;

Best Local Similarity 93.4%; Pred. No. 0; Mismatches 229; Indels 15; Gaps 3;
 Matches 3449; Conservative 0;

Qy	1	TTGACTTCAAAATAGGAAAAATGAGAATGAAATTAATAATGCTTTATCGATTCCAGCTGTA	60
Db	1	TTGACTTCAAAATAGGAAAAATGAGATGAATTAATAATGCTTTATCGATTCCAGCTGTA	60
Qy	61	TCGAATCAATCCACAAATGGATCTATACACGAGTCTCGTATTTAGAGATTCCTTTGTGT	120
Db	61	TCGAATCAATCCGACAAATGAATCTATCAACCGATGCTCGTATTTAGAGATGCTTTGTGT	120
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	121	ATAGCCGAGGGGAACAATATCGATCCATTTGTTAGCGCATCAACAGTCCAAACGGGTATT	180
Qy	181	AACATAGCTGGTAGAATACTAGTGTGTTATAGGGGTACCGTTTGTGTCGACAAATAGCTAGT	240
Db	181	AACATAGCTGGTAGAATACTAGTGTGTTATAGGGGTACCGTTTGTGTCGACAAATAGCTAGT	240
Qy	241	TTTTATAGTTTTCTTTGTTGGTAATATAGCCCGCGGCGAGAGATCAGTGGGAAATTTTC	300
Db	241	TTTTATAGTTTTCTTTGTTGGTAATATAGCCCGCGGCGAGAGATCCTTTGGGAAATTTTC	300
Qy	301	CTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360
Db	301	CTAGAACATGTCGNACAACTTATAGAACACAGTAACAGAAATACTAGGGATACGGCT	360
Qy	361	CTTGCTCGATTACAAGGTTTAGGAGATTCTTTTAGAGCCTTATCAACAGTCACTTGAAGAT	420
Db	361	CTTGCTCGATTACAAGGTTTAGGAAATTCCTTTTAGAGCCTTATCAACAGTCACTTGAAGAT	420
Qy	421	TGGTCAGAAACCGTGATGATGCHAGAACGAGAAGTGTCTTTATACCCATATATAGCC	480
Db	421	TGGTCAGAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCATATATAGCC	480
Qy	481	TTAGAACTTGATTTCTTAATGCGATCCGCTTTTCGCAATTTAGAAACCAAGAGTTTCCA	540
Db	481	TTAGAACTTGATTTCTTAATGCGATCCGCTTTTCGCAATTTAGAAACCAAGAGTTTCCA	540
Qy	541	TTATTAATTTGATATGCTCAAGCTGCAAAATTAACACCTTATTAATTTAGAGATGTCCTCT	600
Db	541	TTATTAATTTGATATGCTCAAGCTGCAAAATTAACACCTTATTAATTTAGAGATGTCCTCT	600
Qy	601	CTTTTTCGTAGTGAATTTGGCTTACATCCGAGAAATTAACCGTTATTTATGAGCGGCAA	660
Db	601	CTTTTTCGTAGTGAATTTGGCTTACATCCGAGAAATTAACCGTTATTTATGAGCGGCAA	660
Qy	661	GTGGAACAAACGAGAGATTATTCGACTATTTCGCTAGAAATGGTATATATACAGGTCATAAT	720

Db	661	GTGGAACAAACGAGAGATTATTCGACTATTTCGCTAGAAATGGTATATATACAGGTCATAAT	720
Qy	721	AGCTTTGAGAGGACAAATCCCAAGTGGTGGCGTTATAAATCAATTCGGTAGAGACTTA	780
Db	721	AAATTTGAGAGGACAAATCTCTGAAAGTGGTGGCGATATAAATCAATTCGGTAGAGACTTA	780
Qy	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA	840
Db	781	ACGCTAGGAGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCTGTTTATCCA	840
Qy	841	ATAAATACCAAGTGTCTCAGTTTAAACAAGGGAAGTTTATACAGAGCGCAATTCGAGCAACAGG	900
Db	841	ATGAATACCAAGTGTCTCAATTAACAGAGAAATTTATACAGATCCCAATTCGGAGAACAAAT	900
Qy	901	GTAAT-----ATGCAAGTATGAATTCGGTATATAAATAAATGAACACTTCGTTTCCGCT	954
Db	901	GCACCTTCAGGATTTGCAAGTACGAATTCGGTTTAAATAAATGACCACTCGTTTCTGCC	954
Qy	955	ATAGAGCTCGCGTTATCCGAGCCCGCATCTACTTGAATTTCTAGAACAACTTACAAT	1014
Db	961	ATAGAGGCTGCGCTTATTAGGCTTCGCACTACTTTGATTTTCCAGAACAGCTTACAAT	1020
Qy	1015	TTTAGCACTTCATCACGATGGAGTGTCTACTAGGCATATGACTTACTTGGCGGGGACACACA	1074
Db	1021	TTACGGGTATTAAAGTCGATGGAGTAATCTCAATATATGAATTTACTGGGTGGGACATAGA	1080
Qy	1075	ATTCAATCTCGGCCAATAGAGCGCGGATTAATACCTCAACCGATGGGTCTACCAATACT	1134
Db	1081	CTTGAATCGGAACAAATAAGGGGGTCAATTAAGTAGTCCCTCGACACACGGAATAACCAATCT	1140
Qy	1135	TCATTAATCTCTGATGATTAATCATCTCTCTCGAGAGCTATATGGACTGAATCATAT	1194
Db	1141	TCATTAATCTCTGATGATTAATCATCTCTCGAGAGCTTTTATAGAACAGAAATCATTT	1200
Qy	1195	GCAGGAGTGTCTTCTATGGGGAATTTACCTTGAACACTTATTCATGGTGTCCCTACTGTAGA	1254
Db	1201	GCAGGAGTAATAAT-----ACTTCTAATCTCTCTGTGAATGGAGTACCTTGGGCTAGA	1254
Qy	1255	TTTAATTTTAGAACCTCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCC	1314
Db	1255	TTTAATTTGAGAAATCCCTGAAATCTCTTT---AGAGGTAGCTTCTCTACTATATAGG	1311
Qy	1315	TATGAGTCACTGGGCTTCAATTTAAAGATTCAGAACTGAAATTAACACAGAAACAAACA	1374
Db	1312	TATCTGAGTGGGGACACAACTATTTGATTCAGAACTGAAATTAACACAGAAACAAACA	1371
Qy	1375	GAACGACCAAAATTAATGAATCATATAGTCTATAGGTATCTCACATAGGGGTCAATTCACAA	1434
Db	1372	GAACGACCAAAATTAATGAATCTTACAGTCAATAGATTATCTAATAATAAGACTAATACAGGA	1431
Qy	1435	TCTAGGGTGCATGTACAGTATATCTTTCGACGACCGTAGTGCAGATCGTACAAATACC	1494
Db	1432	AACACTTTGAGAGCACCGTATATCTTTCGACGACCGTAGTGCAGATCGTACAAATACC	1491
Qy	1495	ATTAGTTTCAGATAGCATAAACAAATACCAATTCGTTAAATTCATTCAACCTTAAATTCAGGT	1554
Db	1492	ATTAGTTTCAGATAGCATAAACAAATACCAATTCGTTAAATTCATTCAACCTTAAATTCAGGT	1551
Qy	1555	ACCTCTGTAGTCAAGTGGCCAGGATTTACAGAGGGGATATATTCGAACTAAACGTTAAT	1614
Db	1552	ACCTCTGTAGTCAAGTGGCCAGGATTTACAGAGGGGATATATTCGAACTAAACGTTAAT	1611
Qy	1615	GGTAGTGTACTAAGTAGTGGGTCTTAATTTTAAATAATACATCATCATACAGGGTATCCGGTG	1674
Db	1612	GGTAGTGTACTAAGTAGTGGGTCTTAATTTTAAATAATACATCATCATACAGGGTATCCGGTG	1671
Qy	1675	AGAGTTTCGTTATGCTCTCTCAACAAATGGTCTCGAGGGTAACTGTCCGAGGAGTACT	1734
Db	1672	AGAGTTTCGTTATGCTCTCTCAACAAATGGTCTCGAGGGTAACTGTCCGAGGAGTACT	1731
Qy	1735	ACTTTTGTATCAAGGATTCCTAGTACTATGAGTGCAAAATGAGTCTTTGACATCTCAATCA	1794

Db	1732	ACTTTTGNATCAAGGATTTCCCTAGTACTATAGAGTGCAAAATGAGTCTTTGGACATCTCAATCA	1791
Qy	1795	TTTAGATTTTGCAGAAATTTCCCTGTAGGTATTTAGTGCATCTCGGCAGTCAAACTGCTGGAATA	1854
Db	1792	TTTAGATTTTGCAGAAATTTCCCTGTAGGTATTTAGTGCATCTCGGCAGTCAAACTGCTGGAATA	1851
Qy	1855	AGTATATAGTAAATATGACAGGTAGACAAACGGTTTCATTTCATTAATAAATTTGAATTCATTTCCA	1914
Db	1852	AGTATATAGTAAATATGACAGGTAGACAAACGGTTTCATTTCATTAATAAATTTGAATTCATTTCCA	1911
Qy	1915	ATTACTGCAACTCTTGCAGACAGAGATACGATTTAGAAAGGCGCAAGAGCGGTGAATGCT	1974
Db	1912	ATTACTGCAACTCTTGCAGACAGAGATATGATTTAGAAAGGCGCAAGAGCGGTGAATGCT	1971
Qy	1975	CTGTTTACTAATACGAATCCAAGAAGATTTGAAAACAGATGTGCACAGATTATCATATTGAT	2034
Db	1972	CTGTTTACTAATACGAATCCAAGAAGGTTCGAAAACAGGTGTGCACAGATTATCATATTGAT	2031
Qy	2035	CAAGTATCCAAATTTAGTGGCGTGTATTCCGATGAATCTTGCTTAGATGAAGAGAGAGAA	2094
Db	2032	GAAGTATCCAAATTTAGTGGCGTGTATTCCGATGAATCTTGCTTAGATGAAGAGAGAGAA	2091
Qy	2095	TTACTTCGAGAAAGTGAATATGCGAACGACTCAGTGTATGATCAAGAGAACTTTACTCCAGAT	2154
Db	2092	TTACTTCGAGAAAGTGAATATGCGAACGACTCAGTGTATGATGAAGAGAACTTTACTCCAGAT	2151
Qy	2155	CCAACTTCACATCCATCAATTAAGCAACACAGACTTCATATCTACTAATGAGCAATCGAAT	2214
Db	2152	CCAACTTCACATCCATCAATTAAGCAACACAGACTTCATATCTACTAATGAGCAATCGAAT	2211
Qy	2215	TTCACTATTCATGAACAAATCTGAACATGGATGGTGGGGAAGTGAGAACTTACAATC	2274
Db	2212	TTCACTATTCATGAACAAATCTGAACATGGATGGTGGGGAAGTGAGAACTTACAATC	2271
Qy	2275	CAGAAAGGAATGACGTATTTAAAGAGAAATTCGTGCACATACGGGGGACTTTAATGAG	2334
Db	2272	CAGAAAGGAATGACGTATTTAAAGAGAAATTCGTGCACATACGGGGGACTTTAATGAG	2331
Qy	2335	TGTTATCCGAGCTATTTATATCAAAAAATAGGAGAGTCCGAAATTAAGAGCTTATACTCGC	2394
Db	2332	TGTTATCCGAGCTATTTATATCAAAAAATAGGAGAGTCCGAAATTAAGAGCTTATACTCGC	2391
Qy	2395	TACCAATTAAGAGGGTATATTTGAAGATPAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2454
Db	2392	TACCAATTAAGTGGCTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2451
Qy	2455	AATCGGAAACATGAAACATTTGGATGGTTCAGGTACCGAGTCCGATAGCGCGCTTCAGTT	2514
Db	2452	AATCGGAAACATGAAACATTTGGATGGTTCAGGTACCGAGTCCGATAGCGCGCTTCAGTT	2511
Qy	2515	GAAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATGCCACACACATTTTGAATGGAAT	2574
Db	2512	GAAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATGCCACACACATTTTGAATGGAAT	2571
Qy	2575	CCTGATCTAGATGTTCTCGCAGAGATGGAGAAAAATTTGCGCATCATTTCCCATCATTTTC	2634
Db	2572	CCTGATCTAGATGTTCTCGCAGAGATGGAGAAAAATTTGCGCATCATTTCCCATCATTTTC	2631
Qy	2635	TCTTTGGATATTTGATATTTGATGCAAGACTTGGATCGAGTATAGCGGTGCGGTGGTA	2694
Db	2632	TCTTTGGATATTTGATATTTGATGCAAGACTTGGATCGAGTATAGCGGTGCGGTGGTA	2691
Qy	2695	TTCAAGATTAAGACGACGAGGAGGTTCATGCAAGACTAGGGAATCTGGAAATTTATTGAAGAG	2754
Db	2692	TTCAAGATTAAGACGACGAGGAGGTTCATGCAAGACTAGGGAATCTGGAAATTTATTGAAGAG	2751
Qy	2755	AAACCAATTTATAGGAGAGCACTGCTCTGTTGTAAGAGACAGAGAAAAAATCGAGAGAC	2814
Db	2752	AAACCAATTTATAGGAGAGCACTGCTCTGTTGTAAGAGACAGAGAAAAAATCGAGAGAC	2811
Qy	2815	AAACGTGAAAACTCAATTTGAAAAAATAAAGAGATATATACAGAGGCAAAAGAGCTGTG	2874
Db	2812	AAACGTGAAAACTCAATTTGAAAAAATAAAGAGATATATACAGAGGCAAAAGAGCTGTG	2871

QY	2875	GATGCTTTATTTGTAGATTCCTCAATATAATAGAAATTACAGCGGATACAAACATTCGGCATG	29334
DB	2872	GATGCTTTATTTGTAGATTCCTCAATATGATAGATTACAGCGGATACAAACATTCGGCATG	29331
QY	2935	ATTCATCCGCGCAGATAAACTTTGTTTCATCGAATTTTCGAGAGCGTTATCTGTGTCAGAAATTTATCT	2994
DB	2932	ATTCATCCGCGCAGATAAACTTTGTTTCATCGAATTTTCGAGAGCGTTATCTTTTCAGAAATTTATCT	2991
QY	2995	GTTATCCCGGGTGTAAATGCGGAAATTTTTTGAAGAAATTAGAAGGTCGCATTTATCACTGCA	3054
DB	2992	GTTATCCCGGTGTAAATGCGGAAATTTTTTGAAGAAATTAGAAGGTCGCATTTATCACTGCA	3051
QY	3055	ATCTCCCTATACGATCGGAGAAATGTCGTTAAAAATGGTGATTTTAAATATGGATTTAGCA	3114
DB	3052	ATCTCCCTATACGATCGGAGAAATGTCGTTAAAAATGGTGATTTTAAATATGGATTTAGCA	3111
QY	3115	TGCTGGAAATGTAAAAAGGCGATGTAGATGTACAAACAGAGCCATCACCGTTCGTGTCCTTGT	3174
DB	3112	TGCTGGAAATGTAAAAAGGCGATGTAGATGTACAAACAGAGCCATCACCGTTCGTGTCCTTGT	3171
QY	3175	ATCCAGAAATGGGAGCAGAAAGTGTCCAAGCAGTTTCGCGTCTGTGTCGGGGCGTGCTAT	3234
DB	3172	ATCCAGAAATGGGAGCAGAAAGTGTCCAAGCAGTTTCGCGTCTGTGTCGGGGCGTGCTAT	3231
QY	3235	ATCTCTCGTGTACAGCGTCAAAAGAGGATATGGAGAGGGTTGCTTAACGATCCATGAA	3294
DB	3232	ATCTCTCGTGTACAGCGTCAAAAGAGGATATGGAGAGGGTTGCTTAACGATCCATGAA	3291
QY	3295	ATCGAGAAACAATACAGACGAACTAAAAATTTAAAAACTGTGAAGAGGAAAGTGTATCCA	3354
DB	3292	ATCGAGAAACAATACAGACGAACTAAAAATTTAAAAACTGTGAAGAGGAAAGTGTATCCA	3351
QY	3355	ACGGATACACGGAACGTGTAAATGATTTATCTGCAACCAAGGTTACAGCAGTATGTAAATCC	3414
DB	3352	ACGGATACACGGAACGTGTAAATGATTTATCTGCAACCAAGGTTACAGCAGTATGTAAATCC	3411
QY	3415	CGTAATCTCGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTGTTAATACAAA	3474
DB	3412	CGTAATCTCGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTGTTAATACAAA	3471
QY	3475	CCGATTTTGAAGAAGAAACGTTATACAGATGTACGAAGAGATAATCAITTTGTGAATATGAC	3534
DB	3472	CCGATTTTGAAGAAGAAACGTTATACAGATGTACGAAGAGATAATCAITTTGTGAATATGAC	3531
QY	3535	AGAGGGTATGTGAATTTATCCACCCTACCAAGCTGGTTATATGACAAAGAAATTTAGAAATAC	3594
DB	3532	AGAGGGTATGTGAATTTATCCACCCTACCAAGCTGGTTATATGACAAAGAAATTTAGAAATAC	3591
QY	3595	TTCCCGAAACCCGATTAAGGTATCGAATTTGAGATTTGGAGAAACCGAAGCGAAGTTTATTTGTA	3654
DB	3592	TTCCCGAAACCCGATTAAGGTATCGAATTTGAGATTTGGAGAAACCGAAGCGAAGTTTATTTGTA	3651
QY	3655	GACAGCGTGAATTTATCTCTTTATGGAGGAATAG	3687
DB	3652	GACAGCGTGAATTTATCTCTTTATGGAGGAATAG	3684

RESULT 5

I90319			PAT 10-AUG-1998
LOCUS	I90319	DNA	linear
LENGTH	3684 bp		
DEFINITION	Sequence 7 from patent US 5723758.		

ACCESSION	I90319
VERSION	I90319.1
	GI:3410259

KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 Unclassified.
1 (bases 1 to 3684)
PAYNE, J., CUMMINGS, D.A., CANNON, R.J.C., NARVA, K.E. and STELMAN, S.
TITLE Bacillus thuringiensis genes encoding lepidopteran-active toxins
JOURNAL Patent: US 5723758-A 7 03-MAR-1998;
FEATURES Location/Qualifiers

[illegible]

Db	2032		GAAGTATCCAAATTTAGTGGCGTGTTATCGGATGAAATCTCGTTGGATGAAGAGAGAA	2091
Qy	2095	TTACTTTGAGAAAGTGAAATATGCGAAACCGACTCAGTGTATGAAAGAAATCTTACTCCAAGAT	2154	
Db	2092	TTACTTTGAGAAAGTGAAATATGCGAAACCGACTCAGTGTATGAAAGAAATCTTACTCCAAGAT	2151	
Qy	2155	CCAAACTTCACATCCATCAATAAGCAACCGAGCTTCATATCTACTAATAGGCAATCGAAT	2214	
Db	2152	CCAAACTTCACATCCATCAATAAGCAACCGAGCTTCATATCTACTAATAGGCAATCGAAT	2211	
Qy	2215	TTCACTCTATCCATGACAACTCTGACATCGATGGTGGGGAAGTGAGACATTTACAATC	2274	
Db	2212	TTCACTCTATCCATGACAACTCTGACATCGATGGTGGGGAAGTGAGACATTTACAATC	2271	
Qy	2275	CAGGAAGCAATGACGTATTTAAAGAGAAATTCAGTCACTACCGGGGCACTTTTAAATGAG	2334	
Db	2272	CAGGAAGCAATGACGTATTTAAAGAGAAATTCAGTCACTACCGGGGCACTTTTAAATGAG	2331	
Qy	2335	TGTTATCCGACGTATTTATACAAAAATAGGAGAGTCCGAAATTTAAAGCTTATACTCGC	2394	
Db	2332	TGTTATCCGACGTATTTATACAAAAATAGGAGAGTCCGAAATTTAAAGCTTATACTCGC	2391	
Qy	2395	TACCAATTAAGAGGGTATATGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2454	
Db	2392	TACCAATTAAGTGGCTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2451	
Qy	2455	AATCGGAAACATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGTATGCGCGCTTTCAATT	2514	
Db	2452	AATCGGAAACATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGTATGCGCGCTTTCAATT	2511	
Qy	2515	GAAGCCCAATCCGAAGTGCGGAGAACCGAATCGATCGCACACACATTTTGAATGGAAT	2574	
Db	2512	GAAGCCCAATCCGAAGTGCGGAGAACCGAATCGATCGCACACATTTTGAATGGAAT	2571	
Qy	2575	CCTGATCTAGATTGTTCTCGCAGNATGGAGAAAAATGTCCGNCATTTCCCATCATTTTC	2634	
Db	2572	CCTGATCTAGATTGTTCTCGCAGNATGGAGAAAAATGTCCGNCATTTCCCATCATTTTC	2631	
Qy	2635	TCCTTGATATTTGATATTTGGATGCACAGACTTGTGATCAGAGAAATCTAGGCGTGTGGTGGA	2694	
Db	2632	TCCTTGATATTTGATATTTGGATGCATAGACTTGTGATCAGAGAACTTAGGCGTGTGGTGGA	2691	
Qy	2695	TTCAAGATTAAGACGCGAGGAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTGAAGAG	2754	
Db	2692	TTCAAGATTAAGACGCGAGGAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTGAAGAG	2751	
Qy	2755	AAACCATTTTAGAGAAAGCACTGTCTCGTGTGAAGAGCAGAGAGAAAAAATGGAGAGAC	2814	
Db	2752	AAACCATTTTAGAGAAAGCACTGTCTCGTGTGAAGAGCAGAGAGAAAAAATGGAGAGAC	2811	
Qy	2815	AAAGCTGAAAACTTACAAATTTGAAAAACAAACGAGTATATACAGGGCAGAAAGAGCTGTG	2874	
Db	2812	AAAGCTGAAAACTTACAAATTTGAAAAACAAACGAGTATATACAGGGCAGAAAGAGCTGTG	2871	
Qy	2875	GATGCTTTATTTGTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG	2934	
Db	2872	GATGCTTTATTTGTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG	2931	
Qy	2935	ATTCAATCGGCGAGATAAATCTGTTTCATCGAAATTCGAGAGGCTTATCTGTGAGAAATTTCT	2994	
Db	2932	ATTCAATCGGCGAGATAAATCTGTTTCATCGAAATTCGAGAGGCTTATCTTTTCAGAAATTTCT	2991	
Qy	2995	GTTATCCCGGTGTAAATGCGGAATTTTGTGAGAAATTTAGAGGTGCGCATTATCACTGCA	3054	
Db	2992	GTTATCCCGGTGTAAATGCGGAATTTTGTGAGAAATTTAGAGGTGCGCATTATCACTGCA	3051	
Qy	3055	ATCTCCCTATACGATCGCAGAAATGTCTGTTAAAAAATGGTCAATTTAAATAATGGATTAGCA	3114	
Db	3052	ATCTCCCTATACGATCGCAGAAATGTCTGTTAAAAAATGGTCAATTTAAATAATGGATTAGCA	3111	
Qy	3115	TGCTTGGAAATGTAAGGGCATGTAGATGTACAACAGAGCCATCACCGTTCTGTCTTGT	3174	

Db	3112	TCCTGGAAATGTAAAGGGCAATGTAGATGTATCAACAGAGCCATCACCGCTTCTGTCCTTTGTT	3117
Qy	3175	ATCCCAAGATGGGAAGCAGAGAAGTGTCTCAACAGCAGTTTCGCGTCTGTCTCGGGGGCGTGGCTAT	3234
Db	3172	ATCCCAAGATGGGAAGCAGAGAAGTGTCTCAACAGCAGTTTCGCGTCTGTCTCGGGGGCGTGGCTAT	3231
Qy	3235	ATCTCTCGTGTCTCAGCGTTACAAAGAGGGATATGGAGAGGGTTGTGTAAACGATCCATGAA	3294
Db	3232	ATCTCTCGTGTCTCAGCGTTACAAAGAGGGATATGGAGAGGGTTGTGTAAACGATCCATGAA	3291
Qy	3295	ATCGAGAACCAATACAGACGAACCTAATAATTTTAAATACTGTGAAGAGGGAAGTGTATCCA	3354
Db	3292	ATCGAGAACCAATACAGACGAACCTAATAATTTTAAATACTGTGAAGAGGGAAGTGTATCCA	3351
Qy	3355	ACGGATACAGGAACCGTGTAAATGATTAATCTGCACACCAAGGTACAGCAGTATGTAATTC	3414
Db	3352	ACGGATACAGGAACCGTGTAAATGATTAATCTGCACACCAAGGTACAGCAGTATGTAATTC	3411
Qy	3415	CGTAATCTCGATATGAGGATGCAATATGAAGTTGATATCTACAGCATCTGTAAATACAAA	3474
Db	3412	CGTAATCTCGATATGAGGATGCAATATGAAGTTGATATCTACAGCATCTGTAAATACAAA	3471
Qy	3475	CCGACTTATGAAGAGAAACGTTATACAGATGTTACGAAGAGATAATCATTTGTGAATATGAC	3534
Db	3472	CCGACTTATGAAGAGAAACGTTATACAGATGTTACGAAGAGATAATCATTTGTGAATATGAC	3531
Qy	3535	AGAGGGTATGTGAATTTATCCACCACTACCAAGCTGGTTTATATGACAAAGAAATTTAGAAATC	3594
Db	3532	AGAGGGTATGTGAATTTATCCACCACTACCAAGCTGGTTTATATGACAAAGAAATTTAGAAATC	3591
Qy	3595	TTCCAGAAAACCGATTAAGTATCGAATGAGATTTGGAGAAAACGGAAGGAAGTTTATTTGTA	3654
Db	3592	TTCCAGAAAACCGATTAAGTATCGAATGAGATTTGGAGAAAACGGAAGGAAGTTTATTTGTA	3651
Qy	3655	GACAGCGTGAATTAATCTCTTATGGAGGAATAG	3687
Db	3652	GACAGCGTGAATTAATCTCTTATGGAGGAATAG	3684
RESULT 6			
AR359400			
LOCUS	AR359400	3684 bp	DNA
DEFINITION	Sequence 62 from patent US 6593293.	linear	PAT 17-AUG-2003
ACCESSION	AR359400		
VERSION	AR359400.1	GI:33765710	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3684)		
AUTHORS	Baum, J. A., Chu, C.-R., Donovan, W. P., Gilmer, A. J., and Rupar, M. J.		
TITLE	Lepidopteran-active Bacillus thuringiensis .delta.-endotoxin		
JOURNAL	compositions and methods of use		
FEATURES	Patent: US 6593293-A 62 15-JUL-2003;		
source	Location/Qualifiers		
	1. .3684		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
Query Match	88.4%;	Score 3260.2;	DB 6; Length 3684;
Best Local Similarity	93.3%;	Pred. No. 0;	
Matches 3445;	Conservative 0;	Mismatches 233;	Indels 15; Gaps 3;
Qy	1	TTGACTTCAAATAGGAAAAATGAGATGAAATTTATAAATGCTTTTATCGATTCAGCTGTA	60
Db	1	TTGACTTCAAATAGGAAAAATGAGATGAAATTTATAAATGCTTTTATCGATTCAGCTGTA	60
Qy	61	TCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTCAGGATTCCTTTGTGT	120
Db	61	TCGAATCATTTCCGCACAAATGAATCTATCAACCGATGCTCGTATTCAGGATTCCTTTGTGT	120
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTTGTTAGCGGCATCAACAGTCCAAACGGGTATT	180

Db 121 ATAGCGAGGGGAACAATATCGATCCATTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
QY 181 AACATAGCTGGTGAAGATACCTAGGTGATAGCGTACCGTTGCTGGCAAAATAGCTAGT 240
Db 181 AACATAGCTGGTGAAGATACCTAGGTGATAGCGTACCGTTGCTGGCAAAATAGCTAGT 240
QY 241 TTTTATAGTTTTCTGCTGGTGAATATAGCGCCCGCGGCGAGAGATCAGTGGGAATTTTC 300
Db 241 TTTTATAGTTTTCTGCTGGTGAATATAGCGCCCGCGGCGAGAGATCCTTGGGAATTTTC 300
QY 301 CTAGAACATGTGCAACACTTATAAATCAACAAATAACAGAAAAATGCTAGGAATACGSCA 360
Db 301 CTAGAACATGTGCAACACTTATAAGACAAACAAGTAAACAGAAATACTAGGGATACGGCT 360
QY 361 CTGCTCGAATACAGAGTTTAGAGATTCCTTTAGAGCTATCAACAGCTCACTTGAAGAT 420
Db 361 CTGCTCGAATACAGAGTTTAGAGAAATTCCTTTAGAGCTATCAACAGCTCACTTGAAGAT 420
QY 421 TGGCTAGAAAAACCGTGATGATGCAAGAACGAGAGAGTTCCTTATACCCATATATAGCC 480
Db 421 TGGCTAGAAAAACCGTGATGATGCAAGAACGAGAGAGTTCCTTATACCCATATATAGCC 480
QY 481 TTAGAACTTGATTTCTTAATGCGATGCGGCTTTTCGCAATTAGAAAAACCAAGAGTTCCA 540
Db 481 TTAGAACTTGATTTCTTAATGCGATGCGGCTTTTCGCAATTAGAAAAACCAAGAGTTCCA 540
QY 541 TTATTAATAGGTATATGCTCAAGCTGCAAAATTTACACCTATTAATTTAGAGATGCGCTCT 600
Db 541 TTATTAATAGGTATATGCTCAAGCTGCAAAATTTACACCTATTAATTTAGAGATGCGCTCT 600
QY 601 CTTTGTGAGTGAATTTGGGCTTACATCGCAGGAAATTCACAGTTATATAGCGCCAA 660
Db 601 CTTTGTGAGTGAATTTGGGCTTACATCCCAAGAAATTCACAGTTATATAGCGCCAA 660
QY 661 GTGGAACAAACGAGAGATTTCCGACTATTCGCTAGTAATGGTATTAATACAGGTCTAAAT 720
Db 661 GTGGAACAAACGAGAGATTTCTGATTAATGCGCAAGATGGTATTAATACGGGTTTAAAT 720
QY 721 AGCTTGAGAGGGAACAAATGCGCAAGTTGGGTGCTTATTAATCAATTCCTGAGAGATCTA 780
Db 721 AATTGAGAGGGAACAAATGCTGAAAGTTGGTGGGATATTAATCAATTCCTGAGAGATCTA 780
QY 781 ACCTTGGGTTATGAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 840
Db 781 ACCTTGGGTTATGAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGGCTGTTATCCA 840
QY 841 ATAAATACGAGTGTCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTCGAGCAACAGGG 900
Db 841 ATGAATACCAAGTGTCTCAATTAACAGAGAAATTTATACAGATCCAATTTGGGAGAACAAAT 900
QY 901 GTAAAT-----ATGGCAAGTATGAATGGTATTAATTAATTAATGACCTTCGTTTCCGCT 954
Db 901 GCACCTTCAGGATTTGCAAGTACGAATGGTTTAAATGAATGACCACTCGTTTCTGCC 960
QY 955 ATAGAGCTCGGTTATCCGAAGCCGCACTACTTGTATTTCTAGAACAACTTCAATTT 1014
Db 961 ATAGAGCTCGGTTATAGGCTTCGCACTACTTGTATTTCCGAAACAGCTTCAATTT 1020
QY 1015 TTTAGCACTTCATCAGATGGAGTGTCTAGGCAATGACTTACTTGGCGGGGCAACACA 1074
Db 1021 TTCAGCGTATTAAGTCGATGGAGTAACTCAATATATGAATTAATGCTGGTGGGCAATAGA 1080
QY 1075 ATTCAATCTCGGCAATAGAGCGGATTAATACCTCAACCGATGGGTCTACCAATACT 1134
Db 1081 CTTGAATCGGAAACAAATAGGGGTCATTAAGTACTTGACACACGCGAAATACCAATACT 1140
QY 1135 TCTATTAATCTGTAGATTTATCTTCTCTCGAGAGCTATATTTGGACTGAATCATAT 1194
Db 1141 TCTATTAATCTGTAGATTTATCACTCTCGAGAGCTTATTAAGAACAGATCATTT 1200
QY 1195 GCAGGAGTCTTCTATGGGGAATTTACCTTGAACTTATCATGGTGTCCCTACTGTTAGA 1254

Db 1201 GCAGGGATAAATAT-----ACTTCTAACTACTCTGTGTAATGGAGTACCTTGGGCTAGA 1254
QY 1255 TTTAATTTTAGGAACCTCTAGAATATCTTTTGAAGAGGTACTCTAACTATAGTCAACCC 1314
Db 1255 TTTAATTTGGGAATATCCCTGTAATCTCTTT---AGAGGTAGCCTTCTCTATATATAGG 1311
QY 1315 TATGAGTCACTCGGCTTCAATTTAAAGATTTCAAGAACTGAAATTAACACAGAAACAACA 1374
Db 1312 TATACTGGAGTGGGACACAACTATTGTTGATTCAGAACTGAAATTAACACAGAAACAACA 1371
QY 1375 GAACGACCAAAATTAATGAATCATATAGTCATAGGTATCTCACATAGGGCTCATTTCCACA 1434
Db 1372 GAACGACCAAAATTAATGAATCTTACAGTCATAGATTAATCTAATATAGACTAATATCAG 1431
QY 1435 TCTAGGCTGATGTACAGTATATCTTCGAGCGCACCGTAGTCAGATCGTCAAAATACC 1494
Db 1432 AACACTTTGAGAGCACAGTATATCTTCGAGCGCACCGTAGTCAGATCGTCAAAATACC 1491
QY 1495 ATTAGTTCAGATAGCATAAACAAAATACCATTTGGTAAATCATTTCAACCTTAATTCAGGT 1554
Db 1492 ATTAGTTCAGATAGCATAAACAAAATACCATTTGGTAAATCATTTCAACCTTAATTCAGGT 1551
QY 1555 ACCTCTGATGTCAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAACCTAACGTTAAT 1614
Db 1552 ACCTCTGATGTCAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAACCTAACGTTAAT 1611
QY 1615 GGTAGTGTACTAAGTATGGGCTCTTAATTTAATAATACATCATTTACAGCGGTATCGCGTG 1674
Db 1612 GGTAGTGTACTAAGTATGGGCTCTTAATTTAATAATACATCATTTACAGCGGTATCGCGTG 1671
QY 1675 AGAGTTCGTTATGCTGCTTCTCAAAACAATGGTCTCGAGGTAACCTCGGAGGGAGTACT 1734
Db 1672 AGAGTTCGTTATGCTGCTTCTCAAAACAATGGTCTCGAGGTAACCTCGGAGGGAGTACT 1731
QY 1735 ACTTTGATCAAGGATTCCTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCA 1794
Db 1732 ACTTTGATCAAGGATTCCTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCA 1791
QY 1795 TTTAGATTTGAGAAATTTCTGTAGGTATTAAGTGCATCTGGCAGTCAAACTGCTGGAATA 1854
Db 1792 TTTAGATTTGAGAAATTTCTGTAGGTATTAAGTGCATCTGGCAGTCAAACTGCTGGAATA 1851
QY 1855 AGTATAAGTAAATATGCAAGGTAGACAAACGTTTCACTTCGTAATAATTTGAATTCATTTCCA 1914
Db 1852 AGTATAAGTAAATATGCAAGGTAGACAAACGTTTCACTTCGTAATAATTTGAATTCATTTCCA 1911
QY 1915 ATTTACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAAGGCGCAAGAGGGGCTGAATGCT 1974
Db 1912 ATTTACTGCAACCTTCGAAGCAGAAATATGATTTAGAAAAGGCGCAAGAGGGGCTGAATGCT 1971
QY 1975 CTGTTTACTAATACGAATCCAAAGAAATTCGAAACAGATGTGACAGATTAATCATATTGAT 2034
Db 1972 CTGTTTACTAATACGAATCCAAAGAAATTCGAAACAGATGTGACAGATTAATCATATTGAT 2031
QY 2035 CAAGTATCCAATTTAGTGGGCTGTTTATCGGATGAATTTCTGCTTGGATGAAAAGAGAGAA 2094
Db 2032 GAAGTATCCAATTTAGTGGGCTGTTTATCGGATGAATTTCTGCTTGGATGAAAAGAGAGAA 2091
QY 2095 TTACTTCGAAAGTGAATATGCGAAACGACTCAGTGTATGAAAGAACTTTACTTCCAAGAT 2154
Db 2092 TTACTTCGAAAGTGAATATGCGAAACGACTCAGTGTATGAAAGAACTTTACTTCCAAGAT 2151
QY 2155 CCAAACTTCACATCCATCAATAAGCAACCAAGCTTCAGTGTATGAAAGAACTTTACTTCCAAGAT 2214
Db 2152 CCAAACTTCACATCCATCAATAAGCAACCAAGCTTCAGTGTATGAAAGAACTTTACTTCCAAGAT 2211
QY 2215 TTTACATCTATCATGAACCAATCTGAACATGGATGGTGGGAAAGTGAGAACATTACCAATC 2274
Db 2212 TTTACATCTATCATGAACCAATCTGAACATGGATGGTGGGAAAGTGAGAACATTACCAATC 2271
QY 2275 CAGGAAAGGAAATGACGTTATTTAAAGAGAAATTAAGTCCACTACCGGGGCTTTTATGAG 2334
Db 2272 CAGGAAAGGAAATGACGTTATTTAAAGAGAAATTAAGTCCACTACCGGGGCTTTTATGAG 2331

```
QY 2335 TGTTATCCGACGTAATTTATATCAAAAAATAGGAGTCCGAATTAAGAGCTTATACGCG 2394
Db 2332 TGTTATCCGACGTAATTTATATCAAAAAATAGGAGGCGGAATTAAGAGCTTATACGCG 2391
QY 2395 TACCAATTAAGAGGTAATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2454
Db 2392 TACCAATTAAGTGGCTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAC 2451
QY 2455 AATCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTT 2514
Db 2452 AATCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTT 2511
QY 2515 GAAAGCCCAATCGGAAGTGGCGAGAACCGAATCGATCGCACCACTTTTGAATGGAAT 2574
Db 2512 GAAAGCCCAATCGGAAGTGGCGAGAACCGAATCGATCGCACCACTTTTGAATGGAAT 2571
QY 2575 CCTGATCTAGATGTTCTTCGACAGATGGAGAGAAAATGCGGCATCATTTCCCATCATTTTC 2634
Db 2572 CCTGATCTAGATGTTCTTCGACAGATGGAGAGAAAATGCGGCATCATTTCCCATCATTTTC 2631
QY 2635 TCTTTGGATATTGATATTGGATGACAGACTTTCATGAGAAATCTAGGCGTGTGGGTGTA 2694
Db 2632 TCTTTGGATATTGATATTGGATGACAGACTTTCATGAGAAATCTAGGCGTGTGGGTGTA 2691
QY 2695 TTCAAGATTTAAGACGCGAAGGTCTATGCAAGTCTAGGAAATCTGGAATTTTATTCGAAG 2754
Db 2692 TTCAAGATTTAAGACGCGAAGGTCTATGCAAGTCTAGGAAATCTGGAATTTTATTCGAAG 2751
QY 2755 AAACCATTTATTAGGAGAGACTGCTGCTGTGGAAGAGAGAGAGAGAGAGAGAGAGAG 2814
Db 2752 AAACCATTTATTAGGAGAGACTGCTGCTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2811
QY 2815 AAACGTGAAACTCAATGGAACAAACAGATATATACAGAGCGCAAGAGAGAGAGAGAG 2874
Db 2812 AAACGTGAAACTCAATGGAACAAACAGATATATACAGAGCGCAAGAGAGAGAGAGAGAG 2871
QY 2875 GATGCTTTATTGATGATTTCTCAATATATATGATTTACAGCGGATACAAACATTTGCGATG 2934
Db 2872 GATGCTTTATTGATGATTTCTCAATATATGATTTACAGCGGATACAAACATTTGCGATG 2931
QY 2935 ATTCAATCGGCGAGATAAATCTGTTTCATGCAATTCGAGAGGCTTATCTGTCGAGAAATATCT 2994
Db 2932 ATTCAATCGGCGAGATAAATCTGTTTCATGCAATTCGAGAGGCTTATCTTTCGAGAAATATCT 2991
QY 2995 GTTATCCCGGCTGTAATCGCGAATTTTGAAGAAATTTAGAGGTCCGATTTATCACTGCA 3054
Db 2992 GTTATCCCGGCTGTAATCGCGAATTTTGAAGAAATTTAGAGGTCCGATTTATCACTGCA 3051
QY 3055 ATCTCCCTATACGATCGAGAAATGCTCGTTAAAAATGGTGATTTTAAATATGGATTAGCA 3114
Db 3052 ATCTCCCTATACGATCGAGAAATGCTCGTTAAAAATGGTGATTTTAAATATGGATTAGCA 3111
QY 3115 TGCTGGAATGTAAGAGGATATAGATGTACAGAGGATACAGGCTTCTGTCCTGTTGT 3174
Db 3112 TGCTGGAATGTAAGAGGATATAGATGTACAGAGGATACAGGCTTCTGTCCTGTTGT 3171
QY 3175 ATCCAGAAATGGAGAGAGAGAGTGTACAGAGGATTCGCTGCTGTCCGCGGCGTGGCTAT 3234
Db 3172 ATCCAGAAATGGAGAGAGAGAGTGTACAGAGGATTCGCTGCTGTCCGCGGCGTGGCTAT 3231
QY 3235 ATCTCCCTGTCACAGCGGTACAAAGAGGATATGGAGAGGTTGTGTAACGATCCATGAA 3294
Db 3232 ATCTCCCTGTCACAGCGGTACAAAGAGGATATGGAGAGGTTGTGTAACGATCCATGAA 3291
QY 3295 ATCGAGAAATATACAGAGAACTAAAAATTTAAAAATCTGTGAAGAGAGAGAGAGTGTATCCA 3354
Db 3292 ATCGAGAAATATACAGAGAACTAAAAATTTAAAAATCTGTGAAGAGAGAGAGTGTATCCA 3351
QY 3355 ACGGATACAGGAGCTGTAAATGATTTATCTGACACAGAGGTACAGAGAGTGTATTTCC 3414
Db 3352 ACGGATACAGGAGCTGTAAATGATTTATCTGACACCAAGGTACAGAGTGTATTTCC 3411
```

```
QY 3415 CGTAATGCTGGATATAGGATCGATATGAAGTTGATCTACAGCATCTGTTAAATACAAA 3474
Db 3412 CGTAATGCTGGATATAGGATCGATATGAAGTTGATCTACAGCATCTGTTAAATACAAA 3471
QY 3475 CCGACTTATGAAGAGAAACGATATACAGATGTACGAGAGATAATCATTTGGAATATGAC 3534
Db 3472 CCGACTTATGAAGAGAAACGATATACAGATGTACGAGAGATAATCATTTGGAATATGAC 3531
QY 3535 AGAGGGTATGGAATTTATCCACCTACCGCTGCTGTATATGACAAAAAGAAATAGAAATAC 3594
Db 3532 AGAGGGTATGGAATTTATCCACCTACCGCTGCTGTATATGACAAAAAGAAATAGAAATAC 3591
QY 3595 TTCCAGAAACCGGATAGGATTTGAGATTTGAGAAACCGAAAGGAAAGTTTATTGTA 3654
Db 3592 TTCCAGAAACCGGATAGGATTTGAGATTTGAGAAACCGAAAGGAAAGTTTATTGTA 3651
QY 3655 GACAGCGTGGAAATTTACTCTCTTATGGAGGAATAG 3687
Db 3652 GACAGCGTGGAAATTTACTCTCTTATGGAGGAATAG 3684

RESULT 7
AX098692
LOCUS AX098692 Sequence 62 from Patent WO0119859. DNA linear PAT 02-APR-2001
DEFINITION AX098692
ACCESSION AX098692
VERSION AX098692.1 GI:13537934
KEYWORDS
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
REFERENCE 1
AUTHORS Baum, J.A., Chu, C.R., Donovan, W.P., Gilmer, A.J. and Rupar, M.J.
TITLE Lepidopteran-active bacillus thuringiensis delta-endotoxin
JOURNAL compositions and methods of use
PATENT: WO 0119859-A 62 22-MAR-2001;
MONSANTO COMPANY (US)
FEATURES
source Location/Qualifiers
1..3684
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"

ORIGIN
Query Match 88.4%; Score 3260.2; DB 6; Length 3684;
Best Local Similarity 93.3%; Fred. No. 0;
Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

QY 1 TTGACTTCAATAGGAAAAATGAGAATGAAATTAATAATGCTTTTATCGATTCAGCTGTA 60
Db 1 TTGACTTCAATAGGAAAAATGAGAATGAAATTAATAATGCTTTTATCGATTCAGCTGTA 60
QY 61 TCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTTGAGAGATCTTTTGTGT 120
Db 61 TCGAATCATTTCCGACAAAATGAATCTATCAACCGATGCTCGTATTTGAGAGATCTTTGTGT 120
QY 121 ATACCGAGGGGGAATATATCAATCCACTTCTGTAGCGCATCAACAGTCCAAACCGGTATT 180
Db 121 ATACCGAGGGGGAACAATATCGATCCATTTGTTAGCGCATCAACAGTCCAAACCGGTATT 180
QY 181 AACATAGCTGTAGTAATACTAGGTGTTATTAGCGGTACCGTTTGTCTGGACAAATAGCTAGT 240
Db 181 AACATAGCTGTAGTAATACTAGGTGTTATTAGCGGTACCGTTTGTCTGGACAAATAGCTAGT 240
QY 241 TTTTATAGTTTTCTTTGTTGGAATTTATGCGCCCGCGCAGAGATCACTAGTGGGAAATTTTC 300
Db 241 TTTTATAGTTTTCTTTGTTGGAATTTATGCGCCCGCGCAGAGATCCCTTGGGAAATTTTC 300
QY 301 CTAGAACATGTCGAACAATTTATATAACAAATATACAGAAATATGCTAGGAAATACGCA 360
Db 301 CTAGAACATGTCGAACAATTTATATAAGCAACAAGTAACAGAAATATCTAGGATACGGCT 360
```

QY 361 CTTGCTCGATTACAAGTTTACGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
Db 361 CTTGCTCGATTACAAGTTTAGAAATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
QY 421 TGGCTAGAAAAACGGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCAATATATAGCC 480
Db 421 TGGCTAGAAAAACGGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCAATATATAGCC 480
QY 481 TTAGAACTTGATTTCTTAATCGGATGCGGCTTTTCGCAATTTAGAAAAACGAAGTTCCA 540
Db 481 TTAGAACTTGATTTCTTAATCGGATGCGGCTTTTCGCAATTTAGAAAAACGAAGTTCCA 540
QY 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTCAGAGATGCGCTCT 600
Db 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTCAGAGATGCGCTCT 600
QY 601 CTTTTTGGTAGTAATTTGGGCTTACATGCGAGGAAATTCACCGTTATTTATGAGCGCAA 660
Db 601 CTTTTTGGTAGTAATTTGGGCTTACATCCCAAGAAATTCACCGTTATTTATGAGCGCAA 660
QY 661 GTGGAAACAAACGAGAGATTATTTCCGACTATTTGCGTAGAAATGGTATATACAGGTCTAAAT 720
Db 661 GTGGAAACAAACGAGAGATTATTTCCGACTATTTGCGTAGAAATGGTATATACAGGTCTAAAT 720
QY 721 AGCTTGAGAGGACAAATGCCCAAGTTGGGTGGCTTATATCAATTCGCTAGAGATCTA 780
Db 721 AATTTGAGAGGACAAATGCTGAAAGTTGGTTGCGATATANTCAATTCGCTAGAGACTTA 780
QY 781 ACGTTAGGGGTATAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA 840
Db 781 ACGTTAGGGGTATAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA 840
QY 841 ATAAATACGAGTGTCAAGTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGG 900
Db 841 ATGAATACCGTGTCAATTAACAGAGAAATTTATACAGATCCAAATTTGGAGAACAAAT 900
QY 901 GTAAAT-----ATGGCAAGTATGAATGGTATATTAATTAATGACACTTCGTTTTCGGCT 954
Db 901 GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAAATTAATGACCACTGTTTTCGCT 960
QY 955 ATAGAGACTCGGGTTATCCGAAGCCCGCATCTACTGATTTCTGAGCAACTTACAAAT 1014
Db 961 ATAGAGCTCGGTTATTAGCCCTCCGATCTACTTGATTTTCCAGAACGCTTACAAAT 1020
QY 1015 TTTAGCACTTCATCAGATGGAGTGTACTAGGCATATGACTTACTGGGGGGGACACA 1074
Db 1021 TTCAGCGTATTAAGTCGATGGAGTAACTCAATATATGAATTAATGCTGGGTGGACATAGA 1080
QY 1075 ATTCAATCTCGGCCAATAGGAGCGGATTAATACCTCAACGATGGGTCTACCAATCT 1134
Db 1081 CTTGAATCGGCACATTAAGGGGTCAATTAAGTACCTGGACACACCGGAAATACCAATCT 1140
QY 1135 TCTATTAATCCTGTAGATATCATTTCTCTCGAGACGTAATTTGGACTGAATCATAT 1194
Db 1141 TCTATTAATCCTGTAACTTACAGTTTACATCTCGAGACGTTTATAGAACAGATCATTT 1200
QY 1195 GCAGGAGTCTTCTATGGGGAATTTACCTTGAACCTTATCATGGTGCCTACTGTTAGA 1254
Db 1201 GCAGGGATAAATAT-----ACTTCTAACTACTCTGTGAATGGAGTACTTGGGGCTAGA 1254
QY 1255 TTTAATTTTAGAACCTCCAGAACTCTTTTGAAGAGGTACTGCTAACTATAGTCAACCC 1314
Db 1255 TTTAATTTGGAGAAATCCCTCGAATTTCTCT--AGAGGTAGCCTCTCTATACTATAGG 1311
QY 1315 TATGAGTCACTGGGCTTCAATTTAAAGATTCAGAACTGAAATTCACCAGAGAAACAACA 1374
Db 1312 TATCTGGAGTGGGGACACAACCTATTGATTCAGAACTGAATTCACCAGAGAAACAACA 1371
QY 1375 GAACGACCAAAATPATGAATCATATAGTATAGTATCTCACAATAGGGCTCATTTCAACA 1434
Db 1372 GAACGACCAAAATPATGAATCTTACAGTCAATAGTATCTCAATATAGCAATAATATACGA 1431
QY 1435 TCTAGGGTGCATGATACAGATATATTTCTTGGACGCAACCGGTAGTGCAGATCGTACAAATACC 1494

Db 1432 AACACTTTGAGAGCACCAGTATATTCTTGGACGACCGTAGTCAGATCGTACAAATACC 1491
QY 1495 ATTAGTTTCAGATAGCATACACAAATACCAATTCGGTAAAAATCATTCAACTTAATTCAGGT 1554
Db 1492 ATTAGTTTCAGATAGCATACACAAATACCAATTCGGTAAAAATCATTCAACTTAATTCAGGT 1551
QY 1555 ACCTCTGTAGTCAAGTGGCCAGGATTTACAGGAGGGGATATTAATCCGACTACAGTTAAT 1614
Db 1552 ACCTCTGTAGTCAAGTGGCCAGGATTTACAGGAGGGGATATTAATCCGAACTAACAGTTAAT 1611
QY 1615 GGTAGTGTTACTAAAGTGGGTCTTAATTTTAAATACATCATTAACAGCGGTATCGCGTG 1674
Db 1612 GGTAGTGTTACTAAAGTGGGTCTTAATTTTAAATACATCATTAACAGCGGTATCGCGTG 1671
QY 1675 AGAGTTTCGTTATCTCTCAAAACAATGGTCTCGAGGGTAACTGTCCGAGGGAGTACT 1734
Db 1672 AGAGTTTCGTTATGCTCTCTCAAAACAATGGTCTCGAGGGTAACTGTCCGAGGGAGTACT 1731
QY 1735 ACTTTTGATCAAGGATTCCTTAGTACTATGAGTGCAAAATGAGTCTTTGACATCTCAATCA 1794
Db 1732 ACTTTTGATCAAGGATTCCTTAGTACTATGAGTGCAAAATGAGTCTTTGACATCTCAATCA 1791
QY 1795 TTTAGATTTGCAGAAATTTCTGTAGGTATTAGTGCAATCTGGCAGTCAAACTGCTGGAATA 1854
Db 1792 TTTAGATTTGCAGAAATTTCTGTAGGTATTAGTGCAATCTGGCAGTCAAACTGCTGGAATA 1851
QY 1855 AGTATAAGTAATTAATGTCAGGTAGACAAACGTTTCTACTTTTGATAAAAATGAAATTCATTCCA 1914
Db 1852 AGTATAAGTAATTAATGTCAGGTAGACAAACGTTTCTACTTTGATAAAAATGAAATTCATTCCA 1911
QY 1915 ATTACTGCAACTCTTGCAGCAGAAATACGATTTAGAAAAGGGCGCAAGAGCGGTGAATGCT 1974
Db 1912 ATTACTGCAACTCTTGCAGCAGAAATATGATTTAGAAAAGCGCAAGAGCGGTGAATGCT 1971
QY 1975 CTGTTTACTTAATACGAATCCAAAGAGATTGAAAAACAGATGTGACAGATTTATCATATTGAT 2034
Db 1972 CTGTTTACTTAATACGAATCCAAAGAGTTTGAACAGGTGTGACAGATTTATCATATTGAT 2031
QY 2035 CAAGTATCCAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTAGATGAAAAGAGAGAA 2094
Db 2032 CAAGTATCCAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTGGATGAAAAGAGAGAA 2091
QY 2095 TTACTTTGAGAAAGTCAAAATATCGAAACGACTCAGTGATGAAAGAACTTACTTCTCAAGAT 2154
Db 2092 TTACTTTGAGAAAGTCAAAATATCGAAACGACTCAGTGATGAAAGAACTTACTTCTCAAGAT 2151
QY 2155 CCAAACTTCACATCCATCAATTAAGCAACAGACTTTCATATCTACTAATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCACATCCATCAATTAAGCAACAGACTTTCATATCTAATTAAGCAATCGAAT 2211
QY 2215 TTCACATCTATCCATGAACATCTGAAACATGGATGGTGGGGAAGTGAGAAACATTACAATC 2274
Db 2212 TTCACATCTATCCATGAACATCTGAAACATGGATGGTGGGGAAGTGAGAAACATTACAATC 2271
QY 2275 CAGGAAGGAAATGACGTTATTTAAAGAGAAATTAACGTACACCTACCGGGGACTTTTAAATGAG 2334
Db 2272 CAGGAAGGAAATGACGTTATTTAAAGAGAAATTAACGTACACCTACCGGGTACTTTTAAATGAG 2331
QY 2335 TGTATCCGAGCTATTTATATCAAAAAATAGGAGAGTCCGAAATTTAAAGCTTATACTCGC 2394
Db 2332 TGTATCCGAGCTATTTATATCAAAAAATAGGAGAGTCCGAAATTTAAAGCTTATACTCGC 2391
QY 2395 TACCAATTTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTAT 2454
Db 2392 TACCAATTTAAGTGGCTATTTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTAT 2451
QY 2455 AATCGAAAAACATGAAAACATTGGATGTTCCAGGTACCGAGTCCGATTCGCGCTTTTCAGTT 2514
Db 2452 AATCGAAAAACATGAAAACATTGGATGTTCCAGGTACCGAGTCCGATTCGCGCTTTTCAGTT 2511
QY 2515 GAAAGCCCAATCCGAAGGTGCGAGAACCGAATCGATCGCCACCAATTTTGAATGGAAT 2574

```
Db 2512 GAAAGCCCAATCGGAAGTGGCGGAAACCGAATCGATCGCACCAATTTGAAATGGAAT 2571
Qy 2575 CCGTATCTAGATTTGTTCTCGCAGAGATGGAGAAAATGTGGCCTCATCTCCCATCATTTTC 2634
Db 2572 CCGTATCTAGATTTGTTCTCGCAGAGATGGAGAAAATGTGGCCTCATCTCCCATCATTTTC 2631
Qy 2635 TCTTTGGATATGATTTGGATGCAAGACTTGCATGAGAAATCTAGGCGTGTGGGTGTTA 2694
Db 2632 TCTTTGGATATGATTTGGATGCAAGACTTGCATGAGAAATCTAGGCGTGTGGGTGTTA 2691
Qy 2695 TTCAGATTTAGACCCAGGAGGTCTATCCAGACTAGGGAATCTCGAATTTATTGCAAGAG 2754
Db 2692 TTCAGATTTAGACCCAGGAGGTCTATCCAGACTAGGGAATCTCGAATTTATTGCAAGAG 2751
Qy 2755 AAACCATTTATTAGGAGAACACTGTCTCGTGTGAAGAGAGACAGAAAAAATGAGAGAC 2814
Db 2752 AAACCATTTATTAGGAGAACACTGTCTCGTGTGAAGAGAGACAGAAAAAATGAGAGAC 2811
Qy 2815 AAACGTGAAAACTCAATTTGGAACAAAAAGAGTATATACAGAGGCAAAAAGAGCTGTG 2874
Db 2812 AAACGTGAAAACTCAATTTGGAACAAAAAGAGTATATACAGAGGCAAAAAGAGCTGTG 2871
Qy 2875 GATGCTTTATTGTAGATTTCTCAATATATAGATTTACAGCGGATACAAACATTTGGCATG 2934
Db 2872 GATGCTTTATTGTAGATTTCTCAATATATAGATTTACAGCGGATACAAACATTTGGCATG 2931
Qy 2935 ATTCAATCGGCGAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTGTCAAGATTATCT 2994
Db 2932 ATTCAATCGGCGAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTGTCAAGATTATCT 2991
Qy 2995 GTTATCCCGGTGTAAATTCGCGAAATTTTGAAGAAATTTAGAGGTTCGCATTTATCACTGCA 3054
Db 2992 GTTATCCCGGTGTAAATTCGCGAAATTTTGAAGAAATTTAGAGGTTCGCATTTATCACTGCA 3051
Qy 3055 ATCTCCCTATACGATCGCAGAAATGCTGTTAAATTCGATTTTAAATGGAATTTAGAA 3114
Db 3052 ATCTCCCTATACGATCGCAGAAATGCTGTTAAATTCGATTTTAAATGGAATTTAGAA 3111
Qy 3115 TGCTGGAATGTAAAGGCGCATGTAGATGTACAACAGAGCCATCACCGTTCTGCTCTGTT 3174
Db 3112 TGCTGGAATGTAAAGGCGCATGTAGATGTACAACAGAGCCATCACCGTTCTGCTCTGTT 3171
Qy 3175 ATCCAGAAATGGGAAGCAGAAAGTGTCAACAGAGTTCGCGTCTGTCCGGGCGGTGGCTAT 3234
Db 3172 ATCCAGAAATGGGAAGCAGAAAGTGTCAACAGAGTTCGCGTCTGTCCGGGCGGTGGCTAT 3231
Qy 3235 ATCTCCCGTGTACAGCGGTACAAGAGGATATGAGAGGGTGTGTGAACATCATGAA 3294
Db 3232 ATCTCCCGTGTACAGCGGTACAAGAGGATATGAGAGGGTGTGTGAACATCATGAA 3291
Qy 3295 ATCGAGAACATACAGAGCAACTAAAAATTTAAAACTGTGAAGAGAGAAAGTGTATCCA 3354
Db 3292 ATCGAGAACATACAGAGCAACTAAAAATTTAAAACTGTGAAGAGAGAAAGTGTATCCA 3351
Qy 3355 ACGGATACAGGAACGTGTAATGATTATATCTGCACACCAAGGTACAGCAGTATGTAATTC 3414
Db 3352 ACGGATACAGGAACGTGTAATGATTATATCTGCACACCAAGGTACAGCAGTATGTAATTC 3411
Qy 3415 CGTAATGCTGATATGAGATGCAATATGAGTGTATCTACAGCATCTGTTAATTAACAA 3474
Db 3412 CGTAATGCTGATATGAGATGCAATATGAGTGTATCTACAGCATCTGTTAATTAACAA 3471
Qy 3475 CCGACTTATGAAGAGAAACGTATACAGATGTACGAAGAGATTAATCATTTGTAATATGAC 3534
Db 3472 CCGACTTATGAAGAGAAACGTATACAGATGTACGAAGAGATTAATCATTTGTAATATGAC 3531
Qy 3535 AGAGGGTATGGAATTTATCCACCTACCACTGCTGTTATATGACAAAAAGAAATAGAAATAC 3594
Db 3532 AGAGGGTATGGAATTTATCCACCTACCACTGCTGTTATATGACAAAAAGAAATAGAAATAC 3591
Qy 3595 TTCCAGAAACCGATTAAGGTATGATTTGAGATTTGAGAAACCGAAGGGAAGTTATTGTA 3654
Db 3592 TTCCAGAAACCGATTAAGGTATGATTTGAGATTTGAGAAACCGAAGGGAAGTTATTGTA 3651
```

```
Qy 3655 GACAGCGTGGAAATTAATCTCTTATGAGGAATAG 3687
Db 3652 GACAGCGTGGAAATTAATCTCTTATGAGGAATAG 3684

RESULT 8
BTRCYA4
LOCUS B. thuringiensis cryA4 gene for delta-endotoxin.
DEFINITION B. thuringiensis cryA4 gene for delta-endotoxin.
ACCESSION X06711
VERSION X06711.1 GI:40264
KEYWORDS cryA4 gene; crystal protein; delta-endotoxin; endotoxin.
SOURCE Bacillus thuringiensis
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 3687)
AUTHORS Brizzard,B.L. and Whiteley,H.R.
TITLE Nucleotide sequence of an additional crystal protein gene cloned from Bacillus thuringiensis subsp. thuringiensis
JOURNAL Nucleic Acids Res. 16 (6), 2723-2724 (1988)
MEDLINE 88203216
PUBMED 3362680
REFERENCE 2 (bases 1 to 3687)
AUTHORS Whiteley,H.R.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1988) Whiteley H. R., Department of W, Department of Microbiology, University of Washington, Seattle, WA 98195
FEATURES
source
1. .3687
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="HD-2"
/db_xref="taxon:1428"
1. .3687
/note="unnamed protein product; delta-endotoxin (AA 1-1228)"
/codon_start=1
/transl_table=11
/protein_id="CAA29898.1"
/db_xref="GI:580949"
/db_xref="GOA:P05517"
/db_xref="Swiss-Prot:P05517"
/transl_table="MTSNRKNENIINAVSNHSAQMDLLDPARIEDSLICAEGNNDIP FVSASTVQTGINAGRLGVLGVPPAGQLASFYSFLVGLMWRDQWEIPLVEUQL INQITENANTALARQLGSDSPRAYQOQSDLEWLNRRDRLSRVLTQVIALEQDF LNAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEFGLTSQETQYRQVER TRDYSYCVSWYNTGLNSLRGTNAASWVRYNQFRDLTLGLVLDLVALFPSTYDRTYPI NTSALQTRVYTDATGATGVNMAWMNNAPSAIEAAAIKSPHLLDLEQITF SASRWSNTRHMTVWRGHTIQSRPIGGLANTSTHGTNTSINPVLTPASRDVVTES YAGVLLNGIYLEPIHGVTVPFNPNTQNSDRGTANYSQYSPGLQKDETELEPP ETTERPNYSYSHRLSHIGILQSRVNPVSTWTHRSADRTNTIGPNRITQIPWKAS ELPGTTVVRGPGTGGDILRRNTGFTVTVNGELTQRYRIGRYASTVDFDF VSRGTNNRFLRTNWSGDELKYNFRAATPTTPTQIQDIIRTSIQGLSGE YVTDKIBIIPVATFEAYDLKAEQAVNALPTNPRRLKTDVTHIDOVSNLVAC LSDFECLDEKRELLEKVKYAKRLSDERNLQDPNFTSINKOPDFISTNEQSNFTSIHE OSBHGWSNLTIOEGNDYFKNYVTLPGTFNFCYPTLYKQIGSELEKATYQLR GYIEDSQDLEIYLIRYNAKHTLDVPGTESLWPLSVESPIGRCEBPNRCAPHEWNPD LDCSRDGEKCAHSHHFSLDIDVGCTDLHENLGVWVFKIKTQEGHARLGNLBFIE KPLGLEALSVRKRAKKRDKREKLOLETRVYTEAKEADVALFVDSQYDLQADTNI GMTAAADLVHRIEAYLSLPGVNASIFELEGLHITAIISLDARNVYVKGDFN NGLTWNVKGHDVQOQSHRSLDVIPEWAEVSOAVRCPGCGVILRVTAKVKGEGE CVTHIELNNDELKFNREBEVYPTDGTCTNDYTAHQGTAGCAGACNSNAGVEDA YEVDTTASGVNKPITYEETVTDVRDHNCHYDRGVNYPVPVAGYVTKLELFFPETDT VWIEIGTEGKFIVDSVELLUMEB"
```

ORIGIN

```
Query Match 85.1%; Score 3136.8; DB 1; Length 3687;
Best Local Similarity 91.2%; Pred. No. 0;
Matches 3375; Conservative 0; Mismatches 297; Indels 30; Gaps 3;
Qy 1 TTGACTTCAATAGAAAAATGAGATGAATTAATTAATGCTTTATCGATTCACAGCTGTA 60
```

Db 1 TTTGACTTCAAAATAGAAAAATAGAAATGAAATATATAAAT-----GCTGTA 45
QY 61 TCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTCAGGATTCCTTTGTC 120
Db 46 TCGAATCATTTCCGACAAATGGATCTATTAACAGATGCTCGTATTCAGGATAGCTTTGTC 105
QY 121 ATAGCCGAGGGGAATATATCAATCACTTTAGGCGCATCAACAGTCCAAACGGGTATT 180
Db 106 ATAGCCGAGGGGAACAATATCGATCCATTTGTTAGCGCATCAACAGTCCAAACGGGTATT 165
QY 181 AACATAGCTGGTGAATATCAATCACTTTAGGCGCATCAACAGTCCAAACGGGTATT 180
Db 166 AACATAGCTGGTGAATATCAATCACTTTAGGCGCATCAACAGTCCAAACGGGTATT 240
QY 241 TTTTATAGTTTCTTCTGGTGAATATGCGCCCGCGGAGATCACTGGGGAATTTTC 300
Db 226 TTTTATAGTTTCTTCTGGTGAATATGCGCCCGCGGAGATCACTGGGGAATTTTC 285
QY 301 CTAGAACATGTGAAACAACTTAAATCAACAAATAACAGAAATGCTTAGGAATACGGCA 360
Db 286 CTAGAACATGTGAAACAACTTAAATCAACAAATAACAGAAATGCTTAGGAATACGGCT 345
QY 361 CTGCTCGAATTAACAGTTTAGAGATTCCTTTAGAGCTTATCAACAGTCACTTTGAAGAT 420
Db 346 CTGCTCGAATTAACAGTTTAGAGATTCCTTTAGAGCTTATCAACAGTCACTTTGAAGAT 405
QY 421 TGGCTAGAAAAACCGTGATGATCAAGAACGAGAGTGTCTTTATACCCATATATAGCC 480
Db 406 TGGCTAGAAAAACCGTGATGATCAAGAACGAGAGTGTCTTTATACCCATATATAGCT 465
QY 481 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTTCGCAATTAGAAACCAAGAGTTCCA 540
Db 466 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTTCGCAATTAGAAACCAAGAGTTCCA 525
QY 541 TTATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTAATTTAGAGATGCTCT 600
Db 526 TTATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTAATTTAGAGATGCTCT 585
QY 601 CTTTTTGTAGTGAATTTGGCTTACATCGCAGGAAATTCACCGTTATTAATGAGAGTCTCT 660
Db 586 CTTTTTGTAGTGAATTTGGCTTACATCGCAGGAAATTCACCGTTATTAATGAGAGTCTCT 645
QY 661 GTGGAACAAACGAGAGATTTCCGACTATTCGAGAAATTCACCGTTATTAATGAGAGTCTCT 720
Db 646 GTGGAACAAACGAGAGATTTCCGACTATTCGAGAAATTCACCGTTATTAATGAGAGTCTCT 705
QY 721 AGCTTGAGAGGACAAATGCGCAAGTTGGGTGCGTTATTAATCAATTCGTTAGAGATCTA 780
Db 706 AGCTTGAGAGGACAAATGCGCAAGTTGGGTGCGTTATTAATCAATTCGTTAGAGATCTA 765
QY 781 ACGTTAGGGGTATAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 840
Db 766 ACGTTAGGGGTATAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 825
QY 841 ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGAGCAATTCGAGCAACAGG 900
Db 826 ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGAGCAATTCGAGCAACAGG 885
QY 901 GTAAATATGCAAGTATGAATTTGGTATATATAATATGCACTTCGTTTTTCGCTATAGAG 960
Db 886 GTAAATATGCAAGTATGAATTTGGTATATATAATATGCACTTCGTTTTTCGCTATAGAG 945
QY 961 ACTGCGGTATTCGAAAGCCCGCATCTACTTTGATTTTCTAGAACACTTACATTTTATAGC 1020
Db 946 GCTGCGGTATTCGAAAGCCCGCATCTACTTTGATTTTCTAGAACACTTACATTTTATAGC 1005
QY 1021 ACTTCATCAGTGGAGTGTCTAGGCAATATGACTTACTTGGCGGGGGCACACAATTCAA 1080
Db 1006 GCTTCATCAGTGGAGTGTCTAGGCAATATGACTTACTTGGCGGGGGCACACAATTCAA 1065
QY 1081 TCTCGGCCAATAGGAGCGGATTTAAATACCTCAACGCTGGTCTACCAATCTCTATT 1140

Db 1066 TCTCGGCCAATAGGAGCGGGAATTAATACCTCAACGCTAGGCGCTACCAATCTCTATT 1125
QY 1141 AATCCTGTAAAGATTTATCAATCTCTCTCGAGAGCTATATTGGAGCTGAATCATATGAGGA 1200
Db 1126 AATCCTGTAAAGATTTATCAATCTCTCTCGAGAGCTATATTGGAGCTGAATCATATGAGGA 1185
QY 1201 GTGCTTCTATGGGGAATTTACCTTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTAAAT 1260
Db 1186 GTGCTTCTATGGGGAATTTACCTTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTAAAT 1245
QY 1261 TTTAGGAACCTCTCAGAAATCTTTTGAAGAGGCTACTGCTAACTATATAGTCAACCTATGAG 1320
Db 1246 TTTAGGAACCTCTCAGAAATCTTTTGAAGAGGCTACTGCTAACTATATAGTCAACCTATGAG 1305
QY 1321 TCACCTGGGCTTCAATTTAAAGATTCAGAAACTGAAATACCAACCAAGAAACAAACAGAACGA 1380
Db 1306 TCACCTGGGCTTCAATTTAAAGATTCAGAAACTGAAATACCAACCAAGAAACAAACAGAACGA 1365
QY 1381 CCAAAATTAAGAAATCATATAGTTCATAGGTTATCTCATAGAGGCTCATTTCAATCTTAGG 1440
Db 1366 CCAAAATTAAGAAATCTTACAGTTCACAGTTTATCTCATATAGGTTATTTTCAATCCAGG 1425
QY 1441 GTGCATGTACCAAGTATATTTCTTGGACGCAACCTAGTGCAGATCGTACAAATACCAATTAGT 1500
Db 1426 GTGAATGTACCGGTATATTTCTTGGACGCAACCTAGTGCAGATCGTACAAATACCAATTAGT 1485
QY 1501 TCAGATAGCATAAACAAATACCATTTGGTAAATCATTTCAACCTTAAATTCAGGTACCTCT 1560
Db 1486 CCAAAATAGAAATCAACCAATTCCTTGGACGCAACCTAGTGCAGATCGTACAAATACCAATTAGT 1545
QY 1561 GTAGTCAGTGGCCAGGATTTTACAGAGGGGATTAATCCGAACTAAAGCTTAAATGCTAGT 1620
Db 1546 GTTGTAGAGGACCAAGATTTTACTGCTGGGATATTTCTTCGAGAAACCAATCTAGTGGGA 1605
QY 1621 GTACTAAGTATGGGCTTAAATTTTAAATTAATCATCATTTACAGCGGTATCGGTGAGAGTT 1680
Db 1606 TTTGACCGATTAAGAGTAACCTGTTAAACGACCAATTAACACAAAGATATCGTATAGGATTC 1665
QY 1681 CTTTATGCTGCTTCTCAAACTAGTCTGAGGGTAACTGCGGAGGAGTACTACTTTT 1740
Db 1666 CCGTATGCTTCAACTGTAGATTTTGTATTTCTTGTATCACGTGGAGGTACTACTGTAAAT 1725
QY 1741 GATCAAGGATTCCTAGTACTATGATGTCGAAATGAGTCTTTGACATCTCAATCAATTTAGA 1800
Db 1726 AATTTAGATTCCTAGTACTAATGAAACAGTGGAGACGAACCTAAATACGGAATTTTGTG 1785
QY 1801 TTTGAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAA--ACTGCTGGAATAAGT 1857
Db 1786 AGCGTGTCTTTTACTACACCTTTTACTTTTACACAAATTCAGATATTAATTCGAACTCT 1845
QY 1858 ATAAGTAAATAGCAGGTAGACAAACGTTTTCATCTTTGTATAAATTTCAATTTCCAAAT 1917
Db 1846 ATTCAAGGCTTATGGAATGGGAAAGTGTATATAGATAAATTTGAAATTTTCCAGTT 1905
QY 1918 ACTGCAACCTTCGAAGCAGAAATACGATTTAGAAGGCGCAAGAGCGGTGAATGCTCTG 1977
Db 1906 ACTGCAACCTTCGAAGCAGAAATGATTTAGAAGGCGCAAGAGCGGTGAATGCTCTG 1965
QY 1978 TTTTACTAATACGAATCCAGAAATTCGAAACAGATGTGACAGATTTATCATATTTGATCAA 2037
Db 1966 TTTTACTAATACGAATCCAGAAATTCGAAACAGATGTGACAGATTTATCATATTTGATCAA 2025
QY 2038 GTATCCAAATTTAGTGGCGTGTATTCGATGAAATTCGATGAAATTCAGATTTATCATATTTGAT 2097
Db 2026 GTATCCAAATTTAGTGGCGTGTATTCGATGAAATTCGATGAAATTCAGATTTATCATATTTGAT 2085
QY 2098 CTTGAGAAATGAAATATCGGAAACGACTCAGTGTAGAAAGAACTTACTCCAAAGATCCA 2157
Db 2086 CTTGAGAAATGAAATATCGGAAACGACTCAGTGTAGAAAGAACTTACTCCAAAGATCCA 2145
QY 2158 AACTTCACATCCATTAAGCAACCAAGCTTCTCATCTACTTAATGAGCAATCGAATTTTC 2217
Db 2146 AACTTCACATCCATTAAGCAACCAAGCTTCTCATCTACTTAATGAGCAATCGAATTTTC 2205

VYIDKIEIIPVTFATEAEYDLERAQEAUNALFTNPNRRLKTDVTDYHIDQVSNLVAC
LSDFCLBEKRELLSKYAKRLSDERNLQDPNFTSINKQPDFISNEQSNFTSIHE
QSEHWGSENIITQEGNDVFNKENTVTLPGTFNECYPTLYOKIGESLKAHYRYOLR
GYIEDSQLEIYLIRYNKHFSLDVPGTESLWPLSVESPIGRGCEPNRCAHPHFNPD
LDCSRDGEKCAHSHHPSLDIDVGCTDLHENLGVWVVFUKIQTQEGHARLGNLEFEE
KPLLGEALSrvKRAEKWRDKREKLOLETKVYVTEAKENDVALFVDSQYDKLQADTHI
GMIHAADKLVRHIREAYLSELPVPGVNAEIPFEBEGHIIITAILSLYDARNVKNQDNF
NGLTCWNVKGHDVQSHRSRLVPEWEAEVSQAVRCPGCGVILRYATYKEGYGEG
CYTHIEINNDDELKFNKREBEVYPTDGTGNDYTAHQGTGACDACSNNRAGYEDA
YEVDTTASVNTKPYTEETDYDVRDNDHCEYDRGVNVPVPGVYVTKELYPEPDTT
VWIEIGETEGKPIVDSVELLMEE"

Query Match 85.0%; Score 3135.2; DB 1; Length 3687;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3374; Conservative 0; Mismatches 298; Indels 30; Gaps 3;

Qy	1	TTGACTTCAAAATAGGAAAAATGAGAAATGAAATATATAATGCTTTATCGATTCCAGCTGTA	60
Db	1	TTGACTTCAAAATAGGAAAAATGAGAAATGAAATATATAATGCTTTATCGATTCCAGCTGTA	45
Qy	61	TCGAATCATTTCCACAAATGGATCTATCACAGATGCTCGTATTTGAGGATTTCTTTGTGT	120
Db	46	TCGAATCATTTCCGACAAATGGATCTATTTACAGATGCTCGTATTTGAGGATGAGCTTGTGT	105
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTTGTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	106	ATAGCCGAGGGGAACATATCGATCCATTTGTAGCGCATCAACAGTCCAAACGGGTATT	165
Qy	181	AACATAGCTGGTAGAATACATAGTGTATTAAGCGTACCGTTTGTGACAAATAGCTAGT	240
Db	166	AACATAGCTGGTAGAATACATAGGCGTATTTGGGCGTACCGTTTGTGACAAATAGCTAGT	225
Qy	241	TTTTATAGTTTCTTTGTTGGTGAATATGCCCCGGCGGAGAGATCAGTGGGAAATTTTC	300
Db	226	TTTTATAGTTTCTTTGTTGGTGAATATGCCCCGGCGGAGAGATCAGTGGGAAATTTTC	285
Qy	301	CTAGAAATGTCGAAACAACTTATAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360
Db	286	CTAGAAATGTCGAAACAACTTATAATCAACAAATAACAGAAATGCTAGGAATACGGCT	345
Qy	361	CTTGCTCGATTTACAAGGTTTACGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
Db	346	CTTGCTCGATTTACAGGTTTACGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	405
Qy	421	TGGCTAGAAACCGTGATGATCAAGAAACGAGAAAGTGTCTTTATACCAATATATAGCC	480
Db	406	TGGCTAGAAACCGTGATGATCAAGAAACGAGAAAGTGTCTTTATACCAATATATAGCT	465
Qy	481	TTAGAACTTGATTTCTTAATGCGATGCGCGCTTTTCGCAATTAGAAACCAAGAAAGTTCCA	540
Db	466	TTAGAACTTGATTTCTTAATGCGATGCGCGCTTTTCGCAATTAGAAACCAAGAAAGTTCCA	525
Qy	541	TTATTAAATGGTATATGCTCAAGCTGCAAAATTTACCTATTTATTTAGAGATGCTCT	600
Db	526	TTATTGATGGTATATGCTCAAGCTGCAAAATTTACCTATTTATTTAGAGATGCTCT	585
Qy	601	CTTTTTGGTAGTAAATTTGGGCTTACATCGCAGGAAATTCACAGTTATTTATGAGCGCCAA	660
Db	586	CTTTTTGGTAGTAAATTTGGGCTTACATCGCAGGAAATTCACAGTTATTTATGAGCGCCAA	645
Qy	661	GTGGAACAAACGAGATTTATTCGACATTTCCGTTAGATGGTATTAATACAGGCTTAAAT	720
Db	646	GTGGAACGAAACGAGATTTATTCGACATTTCCGTTAGATGGTATTAATACAGGCTTAAAT	705
Qy	721	AGCTTGAGAGGGAACAAATCCGCAAGTGTGGGTGCGTTATAATCAATTCGTTAGAGATCTA	780
Db	706	AGCTTGAGAGGGAACAAATCCGCAAGTGTGGGTGCGTTATAATCAATTCGTTAGAGATCTA	765
Qy	781	ACGTTAGGGGTATTTAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA	840
Db	766	ACGTTAGGGGTATTTAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA	825

Qy	841	ATAAATACAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTCGAGCAACAGGG	900
Db	826	ATAAATACAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTCGAGCAACAGGG	885
Qy	901	GTAATATGCGCAAGTATGAATTTGGTATAATAATAATGACACCTTCGTTTTCGCTATAGAG	960
Db	886	GTAATATGCGCAAGTATGAATTTGGTATAATAATAATGACACCTTCGTTTTCGCTATAGAG	945
Qy	961	ACTGCGGTTATCCGAAGCCGCGATCTACTTTGATTTTCTAGAACAACTTACAAATTTTTC	1020
Db	946	GCTGCGGCTATCCGAAGCCGCGATCTACTTTGATTTTCTAGAACAACTTACAAATTTTTC	1005
Qy	1021	ACTTCATACGATGGAGTGTACTAGGCATATGACTTCTGCGGGGGGACACAAATTC	1080
Db	1006	GCTTCATACGATGGAGTGTACTAGGCATATGACTTCTGCGGGGGGACACAAATTC	1065
Qy	1081	TCTCGGCCAATAGGAGCGGATTAATACCTCAACGCAATGGGTCTACCAATATCTTCTATT	1140
Db	1066	TCTCGGCCAATAGGAGCGGATTAATACCTCAACGCAATGGGTCTACCAATATCTTCTATT	1125
Qy	1141	AATCCTGTAAAGATTATCATTTCTCTCGAGACGTATATTGGACTGAATCATATGCAGGA	1200
Db	1126	AATCCTGTAAAGATTATCATTTCTCTCGAGACGTATATTGGACTGAATCATATGCAGGA	1185
Qy	1201	GTGCTTCTATGGGAATTTTACCTTGAACCTATTTCAATGGTGTCCCTACTGTTAGATTTAAT	1260
Db	1186	GTGCTTCTATGGGAATTTTACCTTGAACCTATTTCAATGGTGTCCCTACTGTTAGATTTAAT	1245
Qy	1261	TTTACGAAACCTTCAGAACTTTTGAAGAGGTACTGCTACTATATAGTCAACCCCTATGAG	1320
Db	1246	TTTACGAAACCTTCAGAACTTTTGAAGAGGTACTGCTACTATATAGTCAACCCCTATGAG	1305
Qy	1321	TCACCTGGGCTTCAATTTAAAGATTTCAGAAATTCAGAAATTCAGAAATTCAGAAATTC	1380
Db	1306	TCACCTGGGCTTCAATTTAAAGATTTCAGAAATTCAGAAATTCAGAAATTCAGAAATTC	1365
Qy	1381	CCAAATTAATGATCATATAGTTCATAGTTTATCTCATAGGGCTCATTTTCAATCTAGG	1440
Db	1366	CCAAATTAATGATCATATAGTTCATAGTTTATCTCATAGGGCTCATTTTCAATCTAGG	1425
Qy	1441	GTGCAATGACCAAGTATATTTTGGACCGCAGTGTAGTCAGATCGTACAAATACCAATGAT	1500
Db	1426	GTGCAATGACCAAGTATATTTTGGACCGCAGTGTAGTCAGATCGTACAAATACCAATGAT	1485
Qy	1501	TCAGATAGCATTAACAAATACCAATTTGGTAAATTCATTCACCTTAATTCAGGTAACCTCT	1560
Db	1486	CCAAATAGAAATCACCCAAATCCCAATGGTAAAGCATCCGAACTTCCTCAAGGTACCACT	1545
Qy	1561	GTAGTCAGTGGCCCGAGGATTTACAGGAGGGGATATAATCCGAACTAACCGTTAATGGTAGT	1620
Db	1546	GTGTTAGAGGACCGAGGATTTTACCTGGTGGGATATCTTTCGAAAGCAATACCTGGTGGGA	1605
Qy	1621	GTACTAAGTATGGGTCTTAATTTTAAATTAATACATCATTTACAGCGGTATCGCGTGAAGT	1680
Db	1606	TTTGGACCGGATAAGAGTAACTGTTTAAACCGACCATTAACAAAGATATCGTATAGGATTC	1665
Qy	1681	CGTTATGCTGCTTCTCAACAAATGGTCTGAGGGTAACTGTCCGAGGGAGTACTACTTTT	1740
Db	1666	CGCTATGCTTCAACTGATAGATTTTGAATTTCTTTGATATCACGTGGAGGTACTACTGTAAT	1725
Qy	1741	GATCAAGGATTCCTCTAGTACTATGAGTCAAAATGAGTCTTTTGACATCTCAATATTTAGA	1800
Db	1726	AAATTTAGATTCCTTACGTAACAATGAAACAGTGGAGACGAACTTAAATACCGGAAATTTGTG	1785
Qy	1801	TTTGCAGAAATTCCTGTAGGTATTTAGTGCATCTGGCGAGTCAA--ACTGCTGGAATAGT	1857
Db	1786	AGAGTCTCTTTTACTACACCTTTTACTTTTACACAAATTCAGATATAATTCGAAACGCTCT	1845
Qy	1858	ATAAGTAAATGACAGGTAGACAAACGTTTTCATCTTTGATAAAATTTGAATTCATTTCCAAT	1917
Db	1846	ATTCAAGGCCCTTAGTGGAAATGGGGAAGTGTATATAGATAAAATTTGAAATATTTCCAGTT	1905
Qy	1918	ACTGCAACCTTCGAAAGCAGNATACGATTTAGAAAGGGCGGACAGGGGTGAATGCTCTG	1977

[illegible]

1731 GTTGTAGAGACCAGGATTACTGCTGGGGATATTTCTTCGAAGAACGAATACTGGTGA 1790
1621 GTACTAAGTATGGCTCTTAATTTTAATAATACATCAATACACGGGTATCGCGTGAGAGTT 1680
1791 TTTGACCGTAAAGAGTACTGTAAACGACCATTAACACAAAGATATCGTATAGATTC 1850
1681 CGTTATGCTCTTCAAAACAATGGCTCTGAGGGTAACGTGCGGAGGAGTACTACTTTT 1740
1851 CGCTATGCTTCAACTGTAGATTGATTTCTGATCTCTTGTATCAGCTGGAGGTACTACTGTAAT 1910
1741 GATCAAGGATTCCTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCATTTAGA 1800
1911 AATTTTAGATTCCTAGTACAAATGAACAGTGAGACGAACACTAAATACGGAAATTTTGTG 1970
1801 TTTGACAGAAATTTCTGTAGTATTAAGTGCATCTGGCAGTCAA--ACTGCTGGAATAAGT 1857
1971 AGACGTGCTTTTACTACACCTTTTACTTTTACAAATTCAGATATTAATTCGAAGCTCT 2030
1858 ATAAGTAATAATGCGAGGTAGACAAACGTTTCTACTTTGATAAAATTTGAATTCATCCAAT 1917
2031 ATTCAAGGCTTTAGTGGAAATGGGAAAGTGTATATAGATAAAATTTGAAATTTATTCAGTT 2090
1918 ACTGCAACCTTCGAAGCAGAAATACGATTTTAGAAGGGCGCAAGCGGTGAATGCTCTG 1977
2091 ACTGCAACCTTCGAAGCAGAAATATGATTTTAGAAGAGCGCAAGCGGTGAATGCTCTG 2150
1978 TTTTACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTCATCAA 2037
2151 TTTTACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTCATCAA 2210
2038 GTATCCAAATTTAGTGGCGTTTATCGGATGAATTTCTGCTTAGATGAAAGAGAGAATTA 2097
2211 GTATCCAAATTTAGTGGCGTTTATCGGATGAATTTCTGCTTAGATGAAAGAGAGAATTA 2270
2098 CTTGAGAAAGTCAATATGCGAAACGACTCAGTGTATGAAAGAACTTACTTCAAGATCCA 2157
2271 CTTGAGAAAGTCAATATGCGAAACGACTCAGTGTATGAAAGAACTTACTTCAAGATCCA 2330
2158 AACTTCACATCCATCAATTAAGCAACAGACTTCTATCTACTTAATGAGCAATCGAATTTTC 2217
2331 AACTTCACATCCATCAATTAAGCAACAGACTTCTATCTACTTAATGAGCAATCGAATTTTC 2390
2218 ACATCTATCCATGAACAATCTGAACATGGATGGTGGGAAGTGAGAACATTAACAATCCAG 2277
2391 ACATCTATCCATGAACAATCTGAACATGGATGGTGGGAAGTGAGAACATTAACAATCCAG 2450
2278 GAAGGAATGAGTATTTAAGAGAAATTAAGTCACTACCTACCGGAGCTTTTAATGAGTGT 2337
2451 GAAGGAATGAGTATTTAAGAGAAATTAAGTCACTACCTACCGGAGCTTTTAATGAGTGT 2510
2338 TATCCGAGCTATTTATATCAAAAATAGGAGTGGGAATTAAGAGCTTATACTCGCTAC 2397
2511 TATCCGAGCTATTTATATCAAAAATAGGAGTGGGAATTAAGAGCTTATACTCGCTAC 2570
2398 CAATTAAGAGGTATATTAAGAGATAGTCAAGATTTTAGAGATATATTTGATTCGTTAAT 2457
2571 CAATTAAGAGGTATATTAAGAGATAGTCAAGATTTTAGAGATATATTTGATTCGTTAAT 2630
2458 GCGAAACATGAACAATTTGATGTTCCAGGTACCGAGTCCGATGCGGAGCTTTCAAGTTGAA 2517
2631 GCGAAACATGAACAATTTGATGTTCCAGGTACCGAGTCCGATGCGGAGCTTTCAAGTTGAA 2690
2518 AGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGACCACTATTTTGAATGGAATCCT 2577
2691 AGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGACCACTATTTTGAATGGAATCCT 2750
2578 GATCTAGATTTGTTCTCGAGAGATGGAGAAAATGTGCGCATCATTTCCCATCATTTCTCT 2637
2751 GATCTAGATTTGTTCTCGAGAGATGGAGAAAATGTGCGCATCATTTCCCATCATTTCTCT 2810
2638 TTGGATATTTGATTTGGATGCAAGATTTGATGCAATCTTAGCGGTGGGTGATATTC 2697
2811 TTGGATATTTGATTTGGATGCAAGATTTGATGCAATCTTAGCGGTGGGTGATATTC 2870

2698 AAGATTAAAGCGCAGGAAGGTCAATCAAGACTTAGGAAATCTGGAATTTATTGAAGAGAAA 2757
2871 AAGATTAAAGCGCAGGAAGGTCAATCAAGACTTAGGAAATCTGGAATTTATTGAAGAGAAA 2930
2758 CCAATTTAGGAGAGCACTGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2817
2931 CCAATTTAGGAGAGCACTGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2990
2818 CGTGAAGAACTTACAAATTTGGAACAAACAGAGTATATACAGAGGCAAAAGAGAGCTGTGAT 2877
2991 CGTGAAGAACTTACAAATTTGGAACAAACAGAGTATATACAGAGGCAAAAGAGAGCTGTGAT 3050
2878 GCTTTATTTAGATTTCTCAATATATAGATTTCAAGCGGATACAAACATTTGGCATGATT 2937
3051 GCTTTATTTAGATTTCTCAATATATAGATTTCAAGCGGATACAAACATTTGGCATGATT 3110
2938 CATCGGCAGATAAATCTTTGTCATCGAATCGAGAGGCTTATCTGTGAGAAATTTATCTGTT 2997
3111 CATCGGCAGATAAATCTTTGTCATCGAATCGAGAGGCTTATCTTTCAAGAAATTTACCTGTT 3170
2998 ATCCCGGTGTAAATCGGAAATTTTGAAGAAATTTAGAGGTCGCAATTTATCACTCAATC 3057
3171 ATCCCGGTGTAAATCGGAAATTTTGAAGAAATTTAGAGGTCGCAATTTATCACTCAATC 3230
3058 TCCCTATACGATCGAGAAATGCTGTTAAAAATGGTGATTTTAAATAATGGATTAGCATGC 3117
3231 TCCTTATACGATCGAGAAATGCTGTTAAAAATGGTGATTTTAAATAATGGATTAGCATGC 3290
3118 TGGAAATGAAAGGCGCATGTAGATGTACAAAGAGCCATCACCGTTCTGTCTGTTTATC 3177
3291 TGGAAATGAAAGGCGCATGTAGATGTACAAAGAGCCATCATCGTTCTGACCTGTTATC 3350
3178 CCAGAAATGGGAAGCAGAGTGTCAAGAGCTTCCGCTCTGTCGGGCGGTGGCTATATC 3237
3351 CCAGAAATGGGAAGCAGAGTGTCAAGAGCTTCCGCTCTGTCGGGCGGTGGCTATATC 3410
3238 CTCCGTGTCAAGCGGTACAAAGAGGATATCGAGAGGTTGTTGTAACCATCCATCAATC 3297
3411 CTTGCTGTCAAGCGGTACAAAGAGGATATCGAGAGGCTCGGTACCAATCCATCAATC 3470
3298 GAGAAACATACAGACGAATTAATAATTTTAAAGCTGTGAAAGAGAGAGTGTATCCAAAG 3357
3471 GAGAAACATACAGACGAATTAATAATTTTAAAGCTGTGAAAGAGAGAGTGTATCCAAAG 3530
3358 GATCAGGAACGTGTATGATTTACTGCAACCAAGGTACAGCTGGATGCGCAGATGCA 3405
3531 GATCAGGAACGTGTATGATTTACTGCAACCAAGGTACAGCTGGATGCGCAGATGCA 3590
3406 TGTAAATCCCCTAATGCTGGATATGAGATGCATATGAAGTTGATCTACAGCATCTGTT 3465
3591 TGTAAATCCCCTAATGCTGGATATGAGATGCATATGAAGTTGATCTACAGCATCTGTT 3650
3466 AATTACAAACCGACTTATGAAGAGAAACGATATACAGATGTACGAAGAGATAATCATTTGT 3525
3651 AATTACAAACCGACTTATGAAGAGAAACGATATACAGATGTACGAAGAGATAATCATTTGT 3710
3526 GAATATGACAGAGGATGTGAATTTTCCACCTACCACTGCTGTTATATGACAAAAAGAA 3585
3711 GAATATGACAGAGGATGTGAATTTTCCACCTACCACTGCTGTTATATGACAAAAAGAA 3770
3586 TTAGAATACTTTCCAGAAACCGATAAGTATGGAATTTGAGATTTGAGAAACCGAAGGGAAG 3645
3771 TTAGAATACTTTCCAGAAACCGATAAGTATGGAATTTGAGATTTGAGAAACCGAAGGGAAG 3830
3646 TTTATTTAGACAGCGTGAATTTACTCTTTATGAGGAATAG 3687
3831 TTTATTTAGATAGCGTGAATTTACTCTCATGGAAGAAATAG 3872

RESULT 11
142103
LOCUS

I42103 4074 bp DNA linear PAT 07-OCT-1997

DEFINITION	Sequence 1 from patent US 5628995.
ACCESSION	I42103
VERSION	I42103.1
KEYWORDS	GI:2467598
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
1 (bases 1 to 4074)	
AUTHORS	Peferoen,M., Janssens,S. and Denolf,P.
TITLE	Control of Oestrinia
JOURNAL	Patent: US 5628995-A 1 13-MAY-1997;
FEATURES	Location/Qualifiers
source	1..4074
	/organism="unknown"
ORIGIN	/mol_type="unassigned DNA"
Query Match	85.0%; Score 3133.6; DB 6; Length 4074;
Best Local Similarity	91.1%; Pred. No. 0;
Matches 3373; Conservative	0; Mismatches 299; Indels 30; Gaps 3;
QY	1 TTGACTTCAANTAGGNAATGAGAAATGAAATTAATAATGCTTTATCGATTCCAGCTGTA 60
Db	186 TTGACTTCAANTAGGNAATGAGAAATGAAATTAATAATGCTTTATCGATTCCAGCTGTA 230
QY	61 TCGAATCATTTCCACACAAATGATCTATCACCAGATGCTCGTATTGAGGATCTTTGTGT 120
Db	231 TCGAATCATTTCCGCACAAATGATCTATCACCAGATGCTCGTATTGAGGATCTTTGTGT 120
QY	121 ATAGCGAGGGGAATAATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATT 180
Db	291 ATAGCGAGGGGAACAATAATTGATCCATTGTGTAGCGCATCAACAGTCCAAACGGGTATT 350
QY	181 AACATAGCTGGTAGAATACTAGGTGATATAGCGGTACCGTTTGTGGAACAATAAGTAGT 240
Db	351 AACATAGCTGGTAGAATACTAGGTGATATAGCGGTACCGTTTGTGGAACAATAAGTAGT 240
QY	241 TTTTATAGTTTCTTGTGGTGAATTAATGCGCCCGCGCAGAGATCAGTGGGAAATTTTC 300
Db	411 TTTTATAGTTTCTTGTGGTGAATTAATGCGCCCGCGCAGAGATCAGTGGGAAATTTTC 470
QY	301 CTAGAACATGTCGAACACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
Db	471 CTAGAACATGTCGAACACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCT 530
QY	361 CTGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTTATCAACAGTCACTTGAAGAT 420
Db	531 CTGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTTATCAACAGTCACTTGAAGAT 420
QY	421 TGGCTAGAAACCGTGATGATGCAAGAACGAGAAAGTGTTCTTATACCAATATATAGCC 480
Db	591 TGGCTAGAAACCGTGATGATGCAAGAACGAGAAAGTGTTCTTATACCAATATATAGCT 650
QY	481 TTAGAACTTGATTTCTTAATGCGATGCGCTTTTTCGAATTAAGAAATGAGAGATGCTCT 600
Db	651 TTAGAACTTGATTTCTTAATGCGATGCGCTTTTTCGAATTAAGAAATGAGAGATGCTCT 600
QY	541 TTATTAATGCTATATGCTCAAGCTGCAATTTACACTTATTAATGAGAGATGCTCT 700
Db	711 TTATTAATGCTATATGCTCAAGCTGCAATTTACACTTATTAATGAGAGATGCTCT 770
QY	601 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCAGGTTATTATGAGCGCCA 660
Db	771 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCAGGTTATTATGAGCGCCA 830
QY	661 GTGGAACAACGAGAGATTATCCGACTATTCGTAAGATGGTATAATAACAGGCTTAAT 720
Db	831 GTGGAACAACGAGAGATTATCCGACTATTCGTAAGATGGTATAATAACAGGCTTAAT 720
QY	721 AGCTTGAGAGGGAACAATCCGCAAGTTCGGTGGCTTATAATCAATTCGGTAGAGATCTA 780
Db	891 AGCTTGAGAGGGAACAATCCGCAAGTTCGGTGGCTTATAATCAATTCGGTAGAGATCTA 950
QY	781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 840
Db	951 ACGTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 1010
QY	841 ATAAATACGAGTGTCTCAGTTTAAACAGGGAGTTTATACAGACCAATTTGGAGCAACAGG 900
Db	1011 ATAAATACGAGTGTCTCAGTTTAAACAGGGAGTTTATACAGACCAATTTGGAGCAACAGG 1070
QY	901 GTAATATGCAAGTATGAATTTGGTATAATAATAATGCACTTCGTTTCCGCTATAGAG 960
Db	1071 GTAATATGCAAGTATGAATTTGGTATAATAATAATGCACTTCGTTTCCGCTATAGAG 1130
QY	961 ACTGCGTTATCCGAAGCCGCACTTACTTGTATTAATAATAATGCACTTCGTTTCCGCTATAGAG 1020
Db	1131 GCTGCGTTATCCGAAGCCGCACTTACTTGTATTAATAATAATGCACTTCGTTTCCGCTATAGAG 1190
QY	1021 ACTTCATCGATGGAGTGTCTACTAGGCATATGACTTACTTGGCGGGGACACAAATTCAA 1080
Db	1191 GCTTCATCGATGGAGTGTCTACTAGGCATATGACTTACTTGGCGGGGACACAAATTCAA 1250
QY	1081 TCTCGGCCAATAGGAGCGGATTAATACCTCAACGATGGGTCTACCAATACTTCTATT 1140
Db	1251 TCTCGGCCAATAGGAGCGGATTAATACCTCAACGATGGGTCTACCAATACTTCTATT 1310
QY	1141 AATCTCTGAAGATTATCATTTCTTCTCGAGACGTATATTGGACTGAATCATATGCGAGA 1200
Db	1311 AATCTCTGAAGATTATCATTTCTTCTCGAGACGTATATTGGACTGAATCATATGCGAGA 1370
QY	1201 GTGCTTCTATGGGGAATTTTACCTTGAACCTATTTCATGGTGCCTTACTGTTAGATTAAAT 1260
Db	1371 GTGCTTCTATGGGGAATTTTACCTTGAACCTATTTCATGGTGCCTTACTGTTAGATTAAAT 1430
QY	1261 TTTTAGAACCTTCAGATACTTTTGAAGAGGTCTGCTAACTATAGTCAACCTTATGAG 1320
Db	1431 TTTTAGAACCTTCAGATACTTTTGAAGAGGTCTGCTAACTATAGTCAACCTTATGAG 1490
QY	1321 TCACCTGGGCTTCAATTAAGATTACAGAACTGAATTTACCACAGAAACAAACAGAACGA 1380
Db	1491 TCACCTGGGCTTCAATTAAGATTACAGAACTGAATTTACCACAGAAACAAACAGAACGA 1550
QY	1381 CCAATTAATGAATCATATAGTCAATGTTTCTCAGATAGGGCTCATTTTCAAACTTAGG 1440
Db	1551 CCAATTAATGAATCATATAGTCAATGTTTCTCAGATAGGGCTCATTTTCAAACTTAGG 1610
QY	1441 GTGCATGATACAGTATATTTTGGACGACCGTATGTCAGATCGTACAAATACCAATTAGT 1500
Db	1611 GTGCATGATACAGTATATTTTGGACGACCGTATGTCAGATCGTACAAATACCAATTAGT 1670
QY	1501 TCAGATAGCATAAACAAATACCAATTCGTAAGTAAATCAATCAACCTTAATTCAGGTACCTCT 1560
Db	1671 CCAATTAATGAATCATATAGTCAATGTTTCTCAGATAGGGCTCATTTTCAAACTTAGG 1730
QY	1561 GTAGTACGTGGCCAGGATTTTACAGGAGGGGATTAATCCGAATCAACGTTAATTTGGTAGT 1620
Db	1731 GTTGTAGAGGACCCAGGATTTTACTGGTGGGATTAATTTTTCGAAGACGAATCTGGTGA 1790
QY	1621 GTACTAGTATGGTCTTAATTTTAAATACATCAATTAACAGGGTATCCGCTGAGAGTT 1680
Db	1791 TTTGGACCGATAAGAGTAATCTGTTAAACGGGACCACTTAACACAAAGATATCGTAGGATTC 1850
QY	1681 CGTTATGCTGCTTCTCAAAACAAATGGTCCCTGAGGGTAACTGTCGAGGAGGAGTACTACTTTT 1740
Db	1851 CGCTATGCTTCAACTGTAGATTTTGAATTTCTTTGTATCAGTGGAGGAGTACTACTGTAAT 1910
QY	1741 GATCAAGGATTCCTTAGTATGATGCAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
Db	1911 AATTTTAGATTCTCAGTACATGAAACAGTGGAGACGAATAAAATACGGAATTTTGTG 1970
QY	1801 TTTTGCAAAATTTCTCTAGGTATTTAGTGCATCTGGCAGTCAA---ACTGCTGGAATAGT 1857
Db	1971 AGACGTGCTTTTACTACACCTTTTACTTTTACAAATTTCAAGATATAATTCGAACGTCT 2030
QY	1858 ATAAGTAATAAGCAGGTAGACAAAGCTTTTCACTTTTGATAAAATTTGAATTTCAATTCACCAATT 1917

FEATURES		Location/Qualifiers	
source	1..4798	/organism="Bacillus thuringiensis serovar entomocidus"	
		/mol_type="genomic DNA"	
		/strain="HD-9"	
		/sub_species="entomocidus"	
		/db_xref="taxon:1436"	
promoter	193..231	/note="BtII promoter"	
promoter	215..243	/note="BtI promoter"	
gene	406..4107	/gene="cryIbA2"	
RBS	406..415		
CDS	421..4107	/gene="cryIbA2"	
		/codon_start=1	
		/transl_table=11	
		/product="delta-endotoxin CryIbA2"	
		/protein_id="AAKS1084.1"	
		/db_xref="GI:13959051"	
		/translation="WTSNRKNENEIINAVSNHSAQMDLLPDARIEDSLCTAEGNNIDP FVSASTVOTGINIAGRLIGLVFPAGQLASFYSFLVGLWPRGRDWEIPLFHVQL INQITENARNTALRLQGLGDSFRAYQOSLEDWLENRDDARTSRVLHTQYIALELDF LNMLFAIRNPVPLLMVYQAANLHLLLRDLSLFGSEFGTSGEIQRYERQVER TRDYSDCVENWNTGLSLRGAASWVRYNQRRDITLGLDLVALFSPYDRTTPI NSQLREVVYTDIAIGATGVNMSMWNNAFSAIEAAAIRSPHLLDFLQLTIF SASRSNTRHMYTRGHTIQSRPIGGLNTSHGTANTSIINPVLRFASRDVRYTES YAGLLWGLIYELTHGVPTVRFNFTNPQNSDRGTANYSQPYESPLQLKDSLELPP ETTERPNVESYSHRLSHIGIILQSRVNVVSVWTHRSADRTNIGPNRLTOIPWKAS ELPGQTVVRGPGTGDILLRNTVGGFGPIRVTVNGPLTQVRIGFVASTVDFF VSRGQTVNPRFLRTWNSDELKYNFVRRAFTTPTFTQIQDIIIRTSIQGLSNGE VYIDKIIIPVATFEAYDYDLRAQEAVALNTNPNRLKTDVTDYHIDQVSNLVC LSDFCDEKRELLKVKYAKRLSDERNLLQDPNFTSINKQPDFISTNEQSNFTSHE QSEHGWGSENITIQEONDYKENVYVTLPGTFNECYPTLYQKIGESSELKAYTRYQLR GYIBDSQLEIYILRYNAKHETLDVPGTESLWPLSVESPIGRGCEPNRCAPHEWNPD LDCSRGKECAHHSHFSLDIDVGCTDLHENLGMVVPKIKTQEGHARLONLEFEE KPLLGEALSVKRAEKWRDKREKLETKRVYTEAKBAVDALFVDSQYDRLOQDNI GMHAADKLVHRIEATLSLSELPVPGVNAEIPFEELEGHIIAISLYDARNVKNQDFN NGLTCMVNKGVHDVQSHHRSDLVPEWEAEVSQAVRVCPGCGYILIRVYAYKEGEG CVTIHETENNDELKFEKREEBEVYPTDTCNDYTAHQGTAGCADCACNSRNAGYBDA YEVDTTASVNYKPTYEETVTDVRRDNHCEYDRGVYVPPVPAGYVTKLEYFPEDTD VWIEIGTEGKGFIVDSVELLMEE"	
repeat_region	4200..4238		
		/rpt_type=inverted	
ORIGIN			
Query Match		85.0%; Score 3133.6; DB 1; Length 4798;	
Best Local Similarity		91.1%; Pred. No. 0;	
Matches 3373; Conservative		0; Mismatches 299; Indels 30; Gaps 3;	
Qy	1	TTGACTTCAATAGGAAATGGAATGAAATTAATAATGCTTTATCGATTCGAGCTGTA	60
Db	421	TTGACTTCAATAGGAAATGGAATGAAATTAATAATGCTTTATCGATTCGAGCTGTA	465
Qy	61	TCGAATCATTCACACAAATGGATCTATACACAGATGCTCGTATTGAGGATCTTTGTGCT	120
Db	466	TCGAATCATTCGACACAAATGGATCTATACACAGATGCTCGTATTGAGGATGCTTTGTG	525
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	526	ATAGCCGAGGGGAACAAATATGATTCATTTGTTAGCGCATCAACAGTCCAAACGGGTATT	585
Qy	181	AACATAGCTGGTAGAATCTAGGTGATTAGGCGTACCGTTTCTGCGACAAATAGCTAGT	240
Db	586	AACATAGCTGGTAGAATCTAGGCGTATTGCGCGTACCCTTTGCTGGACAACTAGCTAGT	645
Qy	241	TTTTATAGTTTCTTGTGTGAATTTATGCCCCCGCGGAGAGATCAGTGGGAAATTTTC	300
Db	646	TTTTATAGTTTCTTGTGTGAATTTATGCCCCCGCGGAGAGATCAGTGGGAAATTTTC	705

1786 CCAAAATATGAATCTTTACAGTCACAGGTTATCTCATATAGGTATATAATTTTACAAATCCAGG 1845
1441 GTGCATGTACCAAGTATATTTCTTGGACGACCGGTAGTCAGATCGGTACAAATACCAATAGT 1500
1846 GTGAATGTACCGGTATATTTCTTGGACGACCGGTAGTCAGATCGGTACAAATACCAATAGT 1905
1501 TCAGATAGCAATACCAATACCAATTTGGTAAATCATTTCAACTTTAAATTTCAAGGTACCTCT 1560
1906 CCAAAATAGAAATCACCCAAATCCAAATGGTAAAAGCATCCGAATTTCTTCAAGGTACCACT 1965
1561 GTAGTCAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAATCAACGTTAAATCGTTAGT 1620
1966 GTTGTAGAGGACCAAGGATTTACTGTGGGGATATTTCTTGAAGAACCAATACCTGGTGA 2025
1621 GTACTAAGTATGGGTCTTAATTTTAAATATACATCATATTACAGCGGTATCGCGTGAGAGTT 1680
2026 TTTGACCGATGAAGATTAAGTAACTGTTAAACCGACCATTAACAAAGATATCGGTATAGGATTC 2085
1681 CGTTATGCTGCTTCTCAACCAATGGTCTGAGGGTAACTGTGCGAGGGAGTACTACTTTT 1740
2086 CGCTATGCTTCAACTGTAGATTTTGATTTCTTTGTTATCACCGTGGAGGTACTACTGTAAT 2145
1741 GATCAAGGATTTCCCTAGTACTATAGTGAATAGTCTTTGACATCTCAATCAATTTAGA 1800
2146 AATTTTAGATTTCTTACGTACCAATGAACAGTGGAGACGAATAAATACGGAATTTTGTG 2205
1801 TTTGACGAATTTCTGTAGGTATTTAGTGCATCTGCGAGTCAA--ACTGCTGGATAGT 1857
2206 AGACGTGCTTTTACTACACCTTTTACTTTTACCAAAATTTCAAGATATTAATTCGAACGTCT 2265
1858 ATAAGTAAATAGTCAGAGTAGACAAACGTTTCACTTTGTATAAAATTTGAATTTCAATTTCAAT 1917
2266 ATTCAAGGCTTTAGTGGAAATGGGAAAGTGTATATAGATATAATTTGAATTTTCCAGTT 2325
1918 ACTGCAACCTTCGAAGCAAGATACGATTTAGAAAGGCGCAAGAGCGGTGAATGCTCTG 1977
2326 ACTGCAACCTTCGAAGCAAGATATGATTTAGAAAGAGCGCAAGAGCGGTGAATGCTCTG 2385
1978 TTTACTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATTGATCA 2037
2386 TTTACTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATTGATCA 2445
2038 GTATCCAAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTAGATGAAAGAGAGAAATTA 2097
2446 GTATCCAAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTAGATGAAAGAGAGAAATTA 2505
2098 CTTGGAAGATGAATATTCGGAACGACTCAGTGTATGAAAGAACTTACTCTCAAGATCCA 2157
2506 CTTGGAAGATGAATATTCGGAACGACTCAGTGTATGAAAGAACTTACTCTCAAGATCCA 2565
2158 AACTTTACATCCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2217
2566 AACTTTACATCCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2625
2218 ACATCTATCCATGAACATCTGAACATCTGAACATCTGAACATCTGAACATCTGAACATCTGA 2277
2626 ACATCTATCCATGAACATCTGAACATCTGAACATCTGAACATCTGAACATCTGAACATCTGA 2685
2278 GAAGAAATGACGATTTTAAAGAGAAATTAAGTCACTACCTACCGGAGCTTTTAAATGAGTGT 2337
2686 GAAGAAATGACGATTTTAAAGAGAAATTAAGTCACTACCTACCGGAGCTTTTAAATGAGTGT 2745
2338 TATCCGAGTATTTATCAAAATAGGAGGTCGGAATTTAAAGCTTATACCTACCTACCTACCT 2397
2746 TATCCGAGTATTTATCAAAATAGGAGGTCGGAATTTAAAGCTTATACCTACCTACCTACCT 2805
2398 CAATTAAGAGGATTAATTAAGAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAAT 2457
2806 CAATTAAGAGGATTAATTAAGAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAAT 2865
2458 GCGAAACATGAAACATTTGGATTTCCAGGATACCGAGTCCGATCCGCTGCTTCAAGTTGAA 2517

2866 GCGAAACATGAAACATTTGGATGTTTCCAGGTACCGAGTCCCTATAGCGCGCTTTTCAAGTTGAA 2925
2518 AGCCCAATCGAAGGTGCGGAGAACCGAATCGATGCGCACCACTTTTGAATGGAATCCT 2577
2926 AGCCCAATCGAAGGTGCGGAGAACCGAATCGATGCGCACCACTTTTGAATGGAATCCT 2985
2578 GATCTAGATTTCTTCTTCAGAGATGGAAGAAATGTTGCGCATCATTTCCCATCATTTCTCT 2637
2986 GATCTAGATTTCTTCTTCAGAGATGGAAGAAATGTTGCGCATCATTTCCCATCATTTCTCT 3045
2638 TTGATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2697
3046 TTGATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3105
2698 AAGATTAAAGCGCAGGAGGTCATGCAAGACTGAGGAATCTGGAATTTTATTTGAAGAGAA 2757
3106 AAGATTAAAGCGCAGGAGGTCATGCAAGACTGAGGAATCTGGAATTTTATTTGAAGAGAA 3165
2758 CCATTTATTAGGAGAGCACTGCTCTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2817
3166 CCATTTATTAGGAGAGCACTGCTCTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3225
2818 CGTGAAGAACTAATTTGGAAGAAACAAACGAGTATATACAGAGGCAAAAGAGTGTGGAT 2877
3226 CGTGAAGAACTAATTTGGAAGAAACAAACGAGTATATACAGAGGCAAAAGAGTGTGGAT 3285
2878 GCTTTATTGCTAGATTTCTCAATATATAGATTTAAGCGGATACAAACATTTGGCATGAT 2937
3286 GCTTTATTGCTAGATTTCTCAATATATAGATTTAAGCGGATACAAACATTTGGCATGAT 3345
2938 CATCGGCGAGATAAATTTGTTTTCATCGAATTCGAGAGGCTTATCTGTGAGATTTATCTGTT 2997
3346 CATCGGCGAGATAAATTTGTTTTCATCGAATTCGAGAGGCTTATCTGTGAGATTTATCTGTT 3405
2998 ATCCCGGCTGTAATTTGCGGAAATTTTGAAGAAATTTAGAGGTCGATTTATCACTGCAATC 3057
3406 ATCCCGGCTGTAATTTGCGGAAATTTTGAAGAAATTTAGAGGTCGATTTATCACTGCAATC 3465
3058 TCCCTATACGATCGAGAAATGTCGTTAAATTTGATTTTAAATTTGATTTTAAATTTGATTTTAA 3117
3466 TCCCTATACGATCGAGAAATGTCGTTAAATTTGATTTTAAATTTGATTTTAAATTTGATTTTAA 3525
3118 TGGATTTGAAAGGCGATGTAGATGTACACAGAGCCATCACCGTTCCTGCTTCTGTTTATC 3177
3526 TGGATTTGAAAGGCGATGTAGATGTACACAGAGCCATCACCGTTCCTGCTTCTGTTTATC 3585
3178 CCAGATGGAAGCAGAAATGTCACAAAGCTGTCGCTCTGTCGCGGCGGTGCTATATC 3237
3586 CCAGATGGAAGCAGAAATGTCACAAAGCTGTCGCTCTGTCGCGGCGGTGCTATATC 3645
3238 CTCCGCTCACAGGTCACAAAGGAGGATTTGAGAGGCTTGTGTAAAGTCCATGAAATC 3297
3646 CTCCGCTCACAGGTCACAAAGGAGGATTTGAGAGGCTTGTGTAAAGTCCATGAAATC 3705
3298 GAGAAACATACAGACCACTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 3357
3706 GAGAAACATACAGACCACTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 3765
3358 GATACAGAAACGTTGATTAATGATTTATCTGCAACCAAGGTACAGC-----AGTA 3405
3766 GATACAGAAACGTTGATTAATGATTTATCTGCAACCAAGGTACAGC-----AGTA 3825
3406 TGTAAATTTCCGTAATGCTGGATGAGGATGCAATGATGATGATGATGATGATGATGATGATGAT 3465
3826 TGTAAATTTCCGTAATGCTGGATGAGGATGCAATGATGATGATGATGATGATGATGATGATGAT 3885
3466 AATTAACAAACGCTTATGAAGAGAAACGTTATACAGATGTACGAAGAGATTAATCATTTGT 3525
3886 AATTAACAAACGCTTATGAAGAGAAACGTTATACAGATGTACGAAGAGATTAATCATTTGT 3945
3526 GAATATGACAGAGGATTTGAAATTTATCCACCACTACCACTGCTGTTATATGACAAAAAGAA 3585
3946 GAATATGACAGAGGATTTGCAATTTATCCACCACTACCACTGCTGTTATATGACAAAAAGAA 4005

QY 3586 TTAGAATCTCTCCAGAAACCGATAAGGTATGATTGAGATTGGAGAAACGGAAG 3645
Db TTAGAATCTCTCCAGAAACCGATAAGGTATGAGATTGGAGAAACGGAAG 4065
QY 3646 TTTATTGTGACAGCGTGGAAATCTCTTATGAGGAGATAG 3687
Db TTTATTGTGATAGCGTGGAAATCTCTCTCATGGAAGATAG 4107

RESULT 13
BACCRYIE 3934 bp DNA linear BCT 25-APR-1994
LOCUS Bacillus thuringiensis crystal protein (cryI E75) gene, complete
DEFINITION cds.
ACCESSION L32020.1 GI:474893
VERSION crystal protein.
KEYWORDS Bacillus thuringiensis
SOURCE Bacillus thuringiensis
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 3934)
AUTHORS Donovan, W.P.
JOURNAL Unpublished (1994)
COMMENT Original source text: Bacillus thuringiensis (strain EG 5847) DNA.
FEATURES
Location/Qualifiers
1..3934
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="EG 5847"
/db_xref="taxon:1428"
67..3756
/gene="cryI E75"
67..3756
/gene="cryI E75"
/function="insect toxin"
/codon_start=1
/transl_table=1
/product="crystal protein"
/protein_id="AAA2344.1"
/db_xref="GI:474894"

/translation="MTSNRQVENEINIALSIPTVSNPSTOMNLSDPARIEDSLCVAEV
NNIDPFVSATVQTGINIAGILGVFPAGQLAFYFLVGLWPSGRDPWEIFLE
HYEQLIKQVTEINTNTAIARLEGLRGYRSYQOALETWLDNRNDARSIIILRYVA
LELDITTAIFLRIRNEEVLPMVYQAANLHLLLDASLFGSEWGMASDVNOYQ
EQRYTEYNHCQVYNTGLNNLRGNAESWLRVNFQFRDLTLGLVLDLVALPSPYDT
RTYPINTSAQLTRIEYTDPIGRTPNAPSASTNWNNAPSPSAEALFRPPLHDF
PEOLTIYSASSRSSTQHMVWVGHRLNFRPFGTGLTSTQGLTNTSINPVTLOFTS
RDVVRTESNAGTWILETPVNGVPHARFNPQNIYERGATTSOPYGVGQIQLDS
ETELPTEPTEPNYESYSHKLSHIGLIIGTLRAPVYSWTHRSADRTNTIGPNRTQI
PLVALHLSGVTVGGPGTGGDILRTNTGTGDIRLNVPLSQRYRVRIRYAST
TDLOFFRINTVNTVNGFSRTMNRGDNLEYSFRTAGSPFPFNLAQSTFTLGAOS
FSNQEVYIDVEFPVPAEVTFEARYDLERAQKAVNALFTSTNPRRLKTDVYHLOVS
NMVACLDEFCLDEKELPEKVKYAKRLSDERLLQDPNFTFISQLSFASIDGQSNF
PSINELSHCWGMSANVTIOEGNDVPKENYVTLPGTFNECYPNIDYKIGISELKYAT
RYOLRGVIEDSQLEILIRYNAKHETLDVPGTDSLWLVSPESIGRCGEENPCARHPF
EWNPDLDCRCDGERCAHSHHPTLDDVGGTDLHENLGVWVVFYKTKQEGYARLGNL
ADTNIGMHAADKLVRIRIERYALSELPLVPGVNAEIFEELEIGHITAMSLYDARNVK
NGDFNGLTCMNKVDVQOQHRRSDLVIPWEAEVSOAVRVCPCRGYILRVAYKE
GVEGCVTHIEINNTDELKFNKREEEVPTDCTCNDYTAHQGTACACNRNAGED
AYEVDVTASVNYKPTVEREYTDVRDHNCEYDRGVNVPFVPPAGYVTVKELEYFETD
TVWIEIGETEGEFLVDSVELLME"

ORIGIN

Query Match 72.9%; Score 2687.2; DB 1; Length 3934;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3088; Conservative 593; Indels 15; Gaps 3;
QY 1 TTGACTTCAATAGGAAATGAGATGAAATATAAATGCTTTATCGATCCAGCTGTA 60
Db 67 TTGACTTCAATAGGAAATGAGATGAAATATAAATGCTTTATCGATCCACGCTA 126

QY 61 TCGAATCATTTCCACACAAATGGATCTATCACCAGATGCTCGTATTCAGGATTCCTTTGTG 120
Db TCGAATCATTTCCACCGCAATGAATCTATCACCAGATGCTCGTATTCAGGATTCCTTTGTG 186
QY 121 ATAGCCGAGGGAATATATCAATCCACTTGTAGGCGATCAACAGTCCAAACGGGTATT 180
Db GTAGCCGAGGTGAACAAATATTGATTCATTTGTTAGCGCATCAACAGTCCAAACGGGTATA 246
QY 181 AACATAGCTGGTGAATACTAGGTGATTAGGCGTACCGCTTTGCTGGACAAATAGTACTG 240
Db AACATAGCTGGTGAATAATTGGGCGTATTAGGTGCGCGTTTGTCTGGACAACTAGTACTG 306
QY 241 TTTTATAGTTTCTTCTGTTGGTGAATATGCGCCCGCGGAGAGATCAGTGGGAAATTTTC 300
Db TTTTATAGTTTCTTCTGTTGGGAAATTATGGCCCTAGTGGCAGAGATCCATGGGAAATTTTC 366
QY 301 CTAGAACATGTGCAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAAATACGGCA 360
Db CTGGAACATGTAGAACAACTTATAGACAAACAAAGTAAACAGAAATATCTAGGAATACGGCT 426
QY 361 CTGTCTCGATTACAGGTTTAGAGATTCTTTAGAGCTATCAACAGTCACTTGAAGAT 420
Db ATTGTCTCGATTAGAGGTTTAGGAAAGGCTATAGATCTTACCAGCAGGCTCTTGAAGAT 486
QY 421 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAACTGTTCTTTATACCAATATATAGCC 480
Db TGGTTAGATAACCGAAATGATGCAAGATCAAGAACATTAATCTTTGAGCGCTATGTTGCT 546
QY 481 TTAGAACTTGATTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAAAGTTCCA 540
Db TTAGAACTTGACATTTACTACTGCTATACCGCTTTTCAGAAATACGAAATGAAAGAGTTCCA 606
QY 541 TTATTAAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCCTCT 600
Db TTATTAAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGAGCGATCC 666
QY 601 CTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCAACTGTTATTATGAGCGCAA 660
Db CTTTTGGTAGTGAATGGGGATGGCATCTTCGATGTTAAACCAATATTACCAGAACAA 726
QY 661 GTGGAACAAACGAGAGATTATTCGACTATTCGCTAGTAATGGTATATATACAGTCTTAAAT 720
Db ATCAGATATACAGAGAAATATTCTAACATTTGCGGTACAATGGTATATATACAGGCTTAAAT 786
QY 721 AGCTTGAGAGGGAACAAATGCGCAAGTTGGGTGCGTTTAAATCAATTCCTGAGAGATCTA 780
Db AACTTAAGAGGGAACAAATGCTGAAAGTTGGTTGCGGTATAATCAATTCCTGAGAGACTTA 846
QY 781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 840
Db ACGTTAGGGGTATTAGATTTAGTAGCCCTATTCCCAGCTATGATATCTCGCACTTATCCA 906
QY 841 ATAAATACGAGTGTCTCAGTTTAAACAGGAAAGTTTATACAGACCAATTTGAGAGCAACAGG 900
Db ATCAATACGAGTGTCTCAGTTTAAACAGAGAAATTTATACAGATCCAATTTGGAGAACAAAT 966
QY 901 GTRAAAT-----ATGGCAAGTATGAATTTGGTATAATAATGCACTTCGTTTCCTCGCT 954
Db GCACCTTCAGGATTTGCAAGTACGAAATTTGGTTTAAATAATGCACTTCGTTTCCTGCTG 1026
QY 955 ATAGAGACTCGGGTTATTCGGAAGCCCGCATCTACTTCTGTTGTTTCTAGAACAACTTCAAT 1014
Db ATAGAGGCTGCCATTTTCAGGCCCTCCGCATCTACTTGTGTTTCCAGAACAACTTCAAT 1086
QY 1015 TTTAGACACTTCATCAGATGGAGTGTCTACTAGGCATATGACTTACTGCGGGGGGACACACA 1074
Db TACAGTGCATCAAGCCGTTGGAGTAGCACTCAACATATGAATATTGTTGGGTGGGACATAG 1146
QY 1075 ATTCAATCTCGGCCAATAGAGGCGGATTAATACTCAACGCTGAGGCTTACCA---AT 1131
Db CTTAACTTCGGCCCAATAGAGGGGACATTAATATCTCAACCAAGGACTTACTTAATAAT 1206

1132 ACTTCTATTAAATCCTGTAAAGATTATCAATCTTTCTCGAGAGCTATATTGGACTGAATCA 1191
1207 ACTTCAATTAATCCTGTAAATTAACAGTTTACGTTCTCGAGAGCTTTATAGAACAGATCA 1266
1192 TATGAGAGAGCTCTTATAGGGGAATTTACCTTTGAACCTATTTCATGGTGTCTCCTACTGTT 1251
1267 AATGAGGAGCAAAATAT-----ACTATTACTACTCTCTGTGAATGGAGTACTTTGGGCT 1320
1252 AGATTATTAATTTTAAAGACCTCAGAAATACCTTTTGAAGAGGTAAGTCTTAACCTATAGTCAA 1311
1321 AGATTATTAATTTTAAAGACCTCAGAAATATTTATGAAGAGGCGCCACTTACCTACAGTCAA 1380
1312 CCTATAGTACCTCGGGCTTCAAATTAAGAAATCAGAAATCAGAAATACCAACAGAAACA 1371
1381 CGGTATCAGGAGTGTGGGATTCAAATTAATTTGATTCAGAAATCAGAAATACCAACAGAAACA 1440
1372 ACAGAACGACCAAAATTAAGAAATCATATAGTCTATAGTGTATCTCATATAGGCTCATTTCA 1431
1441 ACAGAACGACCAAAATTAAGAAATCATATAGTCTATAGTGTATCTCATATAGGCTCATTTCA 1500
1432 CAATCTAGGGTGCATGTACCAATATATCTTGGAGCGCCGTAAGTGCAGATCGTACAAAT 1491
1501 GGAACACTTTGAGAGCACCAGTCTATTCTTGGAGCGCATCGTAGTGCAGATCGTACGAAT 1560
1492 ACCATTAGTTCAGATAGCATAACAACAAATACCAATTTGGTAAATCAATCAACCTTAATTTCA 1551
1561 ACGATTGGACCAAAATAGAAATTAACAACAAATACCAATTTGGTAAAGCACTGAATCTTCATTCA 1620
1552 GGTACCTCTGTAGTCAGTGGGCCAGGATTTACAGAGGGGATATTAATCGGAATCAACGTT 1611
1621 GGTGTACTGTGTGTGGAGGGCCAGGATTTACAGGTGGGGATATCTCTGTAGAACAAAT 1680
1612 AATGGTGTAGTACTAAGTATGGTCTTAATTTTAATTAATATACATCAATCACAGCGTATCGC 1671
1681 ACGGGTACATTTGGAGATATACGATTAATTAATTAATGTGCCATATCCCAAGATATCGC 1740
1672 GTGAGAGTTGTTATGCTGCTTCTCAAAACAATGGTCTGAGGGTAACCTGTGCGAGGGAGT 1731
1741 GTAAGGATTCGTTATGCTTCTACTACAGATTTACAAATTTTCAACAGAAATTAATGGAACC 1800
1732 ACTACTTTTGATCAAGGATTCCTAGTACTATAGTGCMAATAGACTTTTGACATCTCAA 1791
1801 ACTGTTAATATTGGTAAATTTCTCAAGAACTATGAATAGGGGGATAATTTAGAAATATAGA 1860
1792 TCATTTAGATTGGCAGAAATTTCTCTAGTATTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
1861 AGTTTATAGACTGCGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCAGCAATTC 1920
1852 ATAAAGTATAAGTAAATGACAGGTAGACAAACGTTTCTACTTTGATAAAATTTGAATTCATT 1911
1921 ACATTTGGGTGCTCAGAGTTTTTCAAAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1980
1912 CCAATTTAGTCAACCTTGAAGCAGAAATAGTATTTAGAAAGGCGCCAGAGCGGTGAAT 1971
1981 CCAGCAGAGGTAACTATTTGAGCGCAATATGATTTTGAAGAGGACCAAAAGCGCGTGAAT 2040
1972 GCTCTGTTTACTAAATACGAATCCAAAGAGATTGAAAACAGATGTGCACAGATTATCATATT 2031
2041 GCTCTGTTTACTTCAAAATCCAAAGAGATTGAAAACAGATGTGCACAGATTATCATATT 2100
2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTCGGAATGAATTTCTGCTTAGATGAAAAGAGA 2091
2101 GACCAAGTGTCCAAATGTTGGCAGTGTATATCAGATGAATTTTGTCTTGGATGAGAAGCGA 2160
2092 GAATTTACTTGAGAAAGTGAATATCGGAAAACGACTCAGTGTATGAAAGAAAATTTACTCCAA 2151
2161 GAAATTTTGAAGAGTGAATATCGGAAGCGACTCAGTGTATGAAAGAAAATTTACTCCAA 2220
2152 GATCCAAACTTCAATCCATTAATGAAGCAACAGACTTCAATCTACTTAATGAGCAATCG 2211
2221 GATCAAAACTTCAATCTAGTGGGCAATTAAGTTTTCGATCCATCGATGGCAATCA 2280
2212 AATTTTCACTATCTCCATGAAACAAATCTGAAATGATGGTGGGGAAGTGAAGCAATTACA 2271

2281 AACTTCCCTCTATTATTAATGAGCTATCTGAAATCATGATGGTGGGAAGTGCGAATGTTACC 2340
2272 ATCCAGGAAGGAATGAGCTATTTAAAGAGAAATTAAGTCACTACCTACCGGGGACTTTTAAT 2331
2341 ATTCCAGGAAGGAATGAGCTATTTAAAGAGAAATTAAGTCACTACCTACCGGGGACTTTTAAT 2400
2332 GAGTGTATCCGACGCTATTTATATCAAAAAATAGGAGAGTGCGAATTTAAAGCTTATACT 2391
2401 GAGTGTATCCAAATTTATATCAAAAAATAGGAGAGTGCGAATTTAAAGCTTATAGC 2460
2392 CGCTACCAATTAAGAGGCTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGT 2451
2461 CGCTATCAATTAAGAGGCTATTTGAAGATAGTCAAGATTTAGAGATATATTTAAATTCGT 2520
2452 TATATGCGAAACATGAAACATTTGATCTCCAGTACCGAGTCCGATGCGGCGCTTTCA 2511
2521 TACATGCAAAAGCATGAAACATTTGATCTCCAGTACCGATTTCCCTATGGCGCTTTCA 2580
2512 GTTGAAGCCCAATCGGAAGGTGGGGAACCGGAATCGATCGCACCAATTTTGAATGG 2571
2581 GTTGAAGCCCAATCGGAAGGTGGGGAACCGGAATCGATCGCACCAATTTTGAATGG 2640
2572 AATCTGATCTAGATTTGTTCTCGAGAGATGGAGAAAAATGCGGCATCATTTCCCATCAT 2631
2641 AATCTGATCTAGATTTGTTCTCGAGAGATGGAGAAAAATGCGGCATCATTTCCCATCAT 2700
2632 TTCTCTTTGGATATTTGATTTGGATGCGACAGATTTGATGAGATCTAGGCGGTGGGTG 2691
2701 TTCACTTTGGATATTTGATTTGGGTGCGACAGCTTTGCAATGAGAACCTTAGGCGGTGGGTG 2760
2692 GTATTTCAAGATTAAGACGCAAGAGGTCTATGCAAGACTTAGGAAATCTCGAAATTTATGAA 2751
2761 GTATTTCAAGATTAAGACGCAAGAGGTCTATGCAAGATTTAGGAATCTCGAAATTTATCGAA 2820
2752 GAGAAACCATTTATTAGGAGAGCACTGCTCTGTGTGAAGAGAGCAGAGAAAAATTTGGAGA 2811
2821 GAGAAACCATTTATTGGAGAGCACTGCTCTGTGTGAAGAGAGCGGAAAAATTTGGAGA 2880
2812 GACAAACGTGAAACACTCAATTTGGAACAAAAACGAGTATATACAGAGGCAAAAGAGCT 2871
2881 GACAAACGTGAAACACTCAATTTGGAACAAAAACGAGTATATACAGAGGCAAAAGAGCT 2940
2872 GTGATGCTTTATTATTGTAGATTCTCAATATATAGATTCAAGCGGATACAAACATTTGCG 2931
2941 GTGATGCTTTATTTCGTAGATTCTCAATATATGATCAATTTCAAGCGGATACAAACATTTGCG 3000
2932 ATGATTCATGCGGAGATATAAATCTGTTTCATCGAAATTCGAGAGGCTTATCTGTCAAAATTA 2991
3001 ATGATTCATGCGGAGATATAAATCTGTTTCATCGAAATTCGAGAGGCTTATCTTTCAGAAATTA 3060
2992 TCTGTATCCCGGTGTAAATGCGGAATTTTGNAGAAATTTAGAGGTGCGATTTATCACT 3051
3061 CCGTGTATCCCGGTGTAAATGCGGAATTTTGNAGAAATTTAGAGGTGCGATTTATCACT 3120
3052 GCAATCTCCCTATACGATGCGAGAAATGCTCGTAAAAATGCTGATTTTAAATTAATGGAATTA 3111
3121 GCAATGCTTTATACGATGCGAGAAATGCTCGTAAAAATGCTGATTTTAAATTAATGGAATTA 3180
3112 GCATGCTGGAATGTAAAGGGCATGTAGATGTACAAAGAGCCATCAACGTTCTGTCTCT 3171
3181 ACATGTTGGAATGTAAAGGGCATGTAGATGTACAAAGAGCCATCACTGTTCTGACCTT 3240
3172 GTTATCCAGAAATGGGAGCAGAAATGTCAACAGAGTTCCGCTGTGTCGGGGCGTGGC 3231
3241 GTTATCCAGAAATGGGAGCAGAAATGTCAACAGAGTTCCGCTGTGTCGGGGCGTGGC 3300
3232 TATATCTCTCGGTGTACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCAT 3291
3301 TATATCTCTCGGTGTACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCAT 3360
3292 GAAATCGAGAAACAAATACAGACGAACTAAAAATTTAAAAATCTGTGAAGAGAGGAATGTAT 3351

1432 CAATCTAGGCTGATGTACCAAGTATATCTTGGAGCGACCGTAGTGCAGATCGTACAAT 1491
1501 GGAACACATTTTGAGAGCACCACTGCTATCTTGGACGCGATCGTAGTGCAGATCGTACGAAT 1560
1492 ACCATTAGTTAGATAGCAATAACAATAACCATTTGGTAAATCAATCAATCAATCAATCA 1551
1561 ACGATTGGACCAATAGAAATTAACAATAACCATTTGGTAAATCAATCAATCAATCAATCA 1620
1552 GGTACCTCTGTAGTCAAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAACTAACGTT 1611
1621 GGTGTACTGTTTGGAGGGCCAGGATTTACAGTGGGGATATCCTTCGTAGAAACAAAT 1680
1612 AATGTAGTGTACTAAGTATGGCTCTTAATTTTAAATAATAACATCAATCAATCAATCAATCA 1671
1681 ACGGTACATTTGGAGATATACGATTAATTAATTAATGTGCCATTAATCCAAAGATATCGC 1740
1672 GTGAGATTTGTTATGCTGCTTCTCAACAATAGTCTGAGGGTAACTGTGCGAGGGAGT 1731
1741 GTAAGGATTTGTTATGCTTCTACTACAGATTTACAAATTTTTCACGAGAAATTAATGGAACC 1800
1732 ACTACTTTTGTATCAAGGATTCCTAGTACTATGAGTGCAGATGAGTCTTTGACATCTCAA 1791
1801 ACTGTTAATTTGGTAATTTCTCAAGAACTATGAATAGGGGGATTAATTTAGAAATATAGA 1860
1792 TCATTTAGATTTGCGAATTTTCTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
1861 AGTTTGAAGACTGCAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCAGACATTC 1920
1852 ATAAGTATAAGTAATAAGTGCAGGTAGACAAAGCTTTCACTTTGATAAAATGAATCAAT 1911
1921 ACATTTGGGTGCTCAGAGTTTTTCAATCAGGAAGTTTATATAGATAGAGTGCAGATTTGTT 1980
1912 CCAATTTACTGCAACCTTCGAACGAGATACAGATTTAGAAAGGGCCAGAGGCGGTGAAT 1971
1981 CAGCAGAGGTAAACATTTGAGGCGAATATGATTTAGAAAGAGCAAAAGGCGGTGAAT 2040
1972 GCTCTGTTTACTAATACGAATCCAAAGAGATTTGAAACAGATGTGCAGAGATTTATCATATT 2031
2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTTGAAACAGATGTGCAGAGATTTATCATATT 2100
2032 GATCAAGTATCAATTTAGTGGCTGTTATTCGGATGAATTTCTGTTAGATGAAGAGAGA 2091
2101 GACCAAGTGTCCAAATATGTGCGATGTTTATCAGATGAATTTGCTTGGATGAGAAGCGA 2160
2092 GAATTTACTTGAGAAAGTGAATATCGGAAACGACTCAGTGTAGTGAAGAACTTACTCCAA 2151
2161 GAATTTATTTGAGAAAGTGAATATCGGAAAGCGACTCAGTGTAGTGAAGAACTTACTCCAA 2220
2152 GATCCAAACTTCCATCCATCAATAAGCAACAGACTTTCAATCTACTAATGAGCAATCG 2211
2221 GATCCAAACTTCCATCCATCAATAAGCAACAGACTTTCAATCTACTAATGAGCAATCA 2280
2212 AATTTCAATCTATCCATGAACAAATCTGAAACATGATGTTGGGGAGGTGAGAACATTA 2271
2281 AACTTCCCTCTTATTAATGAGCTATCTGAAACATGATGTTGGGGAGGTGCGAATGTATACC 2340
2272 ATCCAGGAGGAAATGAGCTATTTAAAGAGAAATACGTCACACTACCGGGACTTTTAAT 2331
2341 ATTCAGGAGGAAATGAGCTATTTAAAGAGAAATACGTCACACTACCGGGACTTTTAAT 2400
2332 GAGTGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCCGGAATTAAGAGCTTATATCT 2391
2401 GAGTGTATCCAAATTTATATCAAAAAATAGGAGAGTCCGGAATTAAGAGCTTATATCT 2460
2392 CGCTACCAATTAAGAGGTATATTAAGAGATTAAGATTAAGATTTAGAGATATATTTGATCGT 2451
2461 CGCTATCAATTAAGAGGTATATTAAGAGATTAAGATTAAGATTTAGAGATATATTTAATCGT 2520
2452 TATATGCAACATGAACATGAGTCTCCAGGTACCGAGTCCGATGCGGCTTCA 2511
2521 TACAATGCAACATGAACATGAGTCTCCAGGTACCGAGTCCGATGCGGCTTCA 2580

2512 GTTGAAGCCCAATCGAAGGTGGGAGAACCCGAATCGATCGCACCAATTTTGAATGG 2571
2581 GTTGAAGCCCAATCGAAGGTGGGAGAACCCGAATCGATCGCACCAATTTTGAATGG 2640
2572 AATCTGTATCTAGATTTGTTCTGCGAGATGAGAGAAATATGTCGCATCATTTCCCATCAT 2631
2641 AATCTGTATCTAGATTTGTTCTGCGAGATGAGAGAAATATGTCGCATCATTTCCCATCAT 2700
2632 TTCTCTTTGGATATTTGATATTTGGATGACACAGCTTGCATGAGAACTTAGCGGTGGGGTG 2691
2701 TTCACTTTGGATATTTGATGTTGGGTGACACAGCTTGCATGAGAACTTAGCGGTGGGGTG 2760
2692 GTATTTCAAGATTAAGACGACGAGAAAGGTCAATCAAGACTAGGGAATCTCGGAATTTATGAA 2751
2761 GTATTTCAAGATTAAGACGACGAGAAAGGTCAATCAAGATTTAGGAATCTCGGAATTTATCGAA 2820
2752 GAGAAACCATTTATTTAGGAGAAAGCACTGCTCTGTTGAAAGAGACGAGAAATATGAGAGA 2811
2821 GAGAAACCATTTATTTGGAGAAAGCACTGCTCTGTTGAAAGAGACGAGAAATATGAGAGA 2880
2812 GACAAACGCTGAAACACTACAAATTTGAAACAAACGAGTATATACAGAGGCAAAAGAGCT 2871
2881 GACAAACGCTGAAACACTACAAATTTGAAACAAACGAGTATATACAGAGGCAAAAGAGCT 2940
2872 GTGGATGCTTTATTTGTAGATTTCTCAATATATATAGATTAACAGCGGATACAAACATTTGCG 2931
2941 GTGGATGCTTTATTTGTAGATTTCTCAATATATATAGATTAACAGCGGATACAAACATTTGCG 3000
2932 ATGATTTGATGCGGAGATTAATCTGTTTCATCGAATTTGAGAGGCTTATCTGTGAGAAATTA 2991
3001 ATGATTTGATGCGGAGATTAATCTGTTTCATCGAATTTGAGAGGCTTATCTGTTGAGAAATTA 3060
2992 TCTGTTATCTCCGCGGTGAAATGCGGAAATTTTGAAGAAATTTAGAGGTCGCAATTTACT 3051
3061 CTTGTTATCTCCGCGGTGAAATGCGGAAATTTTGAAGAAATTTAGAGGTCGCAATTTACT 3120
3052 GCAATCTCTCTATACGATGCGAGAAATGTCGTTAAAAATGTTGATTTTAAATGATGATTA 3111
3121 GCAATGCTCTTATACGATGCGGAGAAATGTCGTTAAAAATGTTGATTTTAAATGATGATTA 3180
3112 GCATGCTGGATGTAATAAGGCGATGTAGATGTACAAAGAGCCATCAACGTTCTGCTCTT 3171
3181 ACATGTTGGATGTAATAAGGCGATGTAGATGTACAAAGAGCCATCACTGTTCTGACCTT 3240
3172 GTTATCCAGAAATGCGGAGCAGAGATGTCACAGCAGTTCCGCTCTGTCGCGGGCGTGGC 3231
3241 GTTATCCAGAAATGCGGAGCAGAGATGTCACAGCAGTTCCGCTCTGTCGCGGGCGTGGC 3300
3232 TATATCTCTCTGTCACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCAT 3291
3301 TATATCTCTCTGTCACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCAT 3360
3292 GAAATCGAGAACAAATACAGAGCAACTAAAAATTTAAAAATCTGTGAAGAGAGAGAGTGTAT 3351
3361 GAAATCGAGAACAAATACAGAGCAACTAAAAATTTAAAAATCTGTGAAGAGAGAGAGTGTAT 3420
3352 CCAACGAGATACAGAAACGTTGATTTATCTACACCAAGGTACAGAGTGTATGAT 3411
3421 CCAACGAGATACAGAAACGTTGATTTATCTACACCAAGGTACAGAGTGTATGAT 3480
3412 TCCCGTAAATGCTGGATATGAGGATGCAATGAAAGTTGATCTACAGCATCTGTTAAATTTAC 3471
3481 TCCCGTAAATGCTGGATATGAGGATGCAATGAAAGTTGATCTACAGCATCTGTTAAATTTAC 3540
3472 AAACCGATTTATGAAGAGAAACGTTATACAGATGTACGAAGATATATCTGTTGAAATAT 3531
3541 AAACCGATTTATGAAGAGAAACGTTATACAGATGTACGAAGATATATCTGTTGAAATAT 3600
3532 GACAGAGGTATGTGAATTTATCCACCACTACAGCTGTTTATATGACAAAAAGAAATAGAA 3591
3601 GACAGAGGTATGTGAATTTATCCACCACTACAGCTGTTTATATGACAAAAAGAAATAGAA 3660
3592 TACTTCCAGAAACCGATTAAGGTATGAGATTTGAGATTTGAGAAACCGAGGAGGATTTATT 3651

```
|||||
Db 3661 TACTTCCGAGAACAGATACAGTATGGATTGAGATTGGAGAAACGGAAGAAAGTTATT 3720
QY 3652 GTAGACAGCGTGAATTTACTCTCTTATGAGGAATAG 3687
Db 3721 GTAGATAGCGTGAATTTACTCTCTCATGGAAGATAG 3756

RESULT 15
138760
LOCUS I38760 3934 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 3 from patent US 5616319.
ACCESSION I38760
VERSION I38760.1 GI:2083238
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 3934)
Donovan, W.P., Tan, Y., Jany, C.S. and Gonz alez, J.e.M. Jr.
Bacillus thuringiensis cryET5 gene and related plasmids, bacteria
and insecticides
JOURNAL Patent: US 5616319-A 3 01-APR-1997;
FEATURES
source Location/Qualifiers
1..3934
/organism="unknown"
/ mol_type="unassigned DNA"
ORIGIN

Query Match 72.9%; Score 2687.2; DB 6; Length 3934;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

QY 1 TTGACTTCAAAATAGGAAATAGAAATGAAATTAATAATGCTTTATCGATTCCACGCTGA 60
Db 67 TTGACTTCAAAATAGGAAATAGAAATGAAATTAATAATGCTTTATCGATTCCACGCTGA 126
QY 61 TCGAATCATTTCCACAAATGATCTATCACCAGATGCTCGTATGAGGATTCCTTTGTGT 120
Db 127 TCGAATCCTTCCACGCAATGATCTATCACCAGATGCTCGTATGAGGATTCCTTTGTGT 186
QY 121 ATAGCCGAGGGGATATATATCAATCCACTTGTATAGCGCATCAACAGTCCAAACGGGTAT 180
Db 187 GTAGCCGAGGGTGAACAATATATGATCCATTTGTATAGCGCATCAACAGTCCAAACGGGTAT 246
QY 181 AACATAGCTGTAGAACTACTAGGTGATATAGCGGTACCGTTGCTGGACAAATAGCTAGT 240
Db 247 AACATAGCTGTAGAACTACTAGGTGATATAGCGGTACCGTTGCTGGACAAATAGCTAGT 306
QY 241 TTTTATAGTTTCTTGTGTGGAATTAATGCCCCCGGCGGAGAGATCAGTGGGAAATTTTC 300
Db 307 TTTTATAGTTTCTTGTGTGGAATTAATGCCCCCGGCGGAGATCAGTGGGAAATTTTC 366
QY 301 CTAGAACATGTGGAACAACTTATATATCAACAATACAGAAATGCTAGGAATACGGCA 360
Db 367 CTGGAACATGTAGAACAACTTATATAGAACAAAGTAAACAGAAATAGCTAGGAATACGGCT 426
QY 361 CTGCTCGATTAACAAGTTTATGAGATATCTTTAGAGCGCTATCAACAGTCACTTGAAGAT 420
Db 427 ATTGCTCGATTAGAGGCTCTAGGAAGGCTATAGATCTTACACAGAGCTCTTGAAGAT 486
QY 421 TGGCTAGAAAACCGTGTAGTGCAGAACGAGAGTGTCTTTATACCCAAATATATAGCC 480
Db 487 TGGTATAGTAACCGAAATGATGCAAGATCAAGAAGCATTAATCTTGAGCGCTATGTGCT 546
QY 481 TTAGAACTTGATTTCTTAATCGATGCGCTTTTTCGAAATTTAGAAACCAAGAGTTCCTCA 540
Db 547 TTAGAACTTTGACATTACTACTGCTATACCGCTTTTTCAGAAATACGAAATGAAGAAGTTCCTCA 606
QY 541 TTATTAATGGTATATGCTCAAGCTCAAAATTTACCTATTATTTATGAGAGATGCGCTCT 600
Db 607 TTATTAATGGTATATGCTCAAGCTCAAAATTTACACCTATTATTTATGAGAGATGCGCTCT 666
```

```
QY 601 CTTTTTGTAGTGAATTTGGCTTTACATCGCAGGAAATTCACGTTATTATGACGCGCAA 660
Db 667 CTTTTTGTAGTGAATTTGGGATGGCATCTTCGATGTTAACCAATATTATCAAGAACAA 726
QY 661 GTGGAACAAACGAGAGATTTATTCGACTATTTGGTGAATGGTATATATACAGGTCTTAAT 720
Db 727 ATCAGATATACAGAGGAATATTCTAACCATTTGGGTACAATGGTATATATACAGGGCTAAAT 786
QY 721 AGCTTTGAGGGACAAATGCGCAAGTTGGGTGCGTTATTAATCAATTCGTTAGAGATCTTA 780
Db 787 AACTTAAGAGGGACAAATGCTGAAGTTGGTTCGGGTATTAATCAATTCGTTAGAGACCTTA 846
QY 781 ACCTTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 840
Db 847 ACCTTTAGGGGTATTAGATTTAGTGGCTATTCCCAAGCTATGATACTCGCACTTATCCA 906
QY 841 ATAAATACGAGTCTCTCAGTTTAAACAGGGAGTTTATACAGAGCGCAATTTGGAGCAACAGG 900
Db 907 ATCAATACGAGTCTCTCAGTTTAAACAGAGAAATTTATACAGATCCCAATTTGGGAGAACAAAT 966
QY 901 GTAAAT-----ATGCGCAAGTATGAATTTGGTATATAATAATGACACTTCGTTTCCGCT 954
Db 967 GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAATAATAATGACCACTCGTTTCTGCC 1026
QY 955 ATAGAGACTGCGGTTTATCCGAAGCCCGCATCTACTTGTGATTTCTAGAACAACTTACAAT 1014
Db 1027 ATAGAGGCTGCCATTTTCAGGCTCCGCATCTACTTGTGATTTTCCAGAACAACTTACAAT 1086
QY 1015 TTTAGCACTTCATCAGATGAGTCTACTAGGCAATGACTTACTTATGCGGGGGGCACACA 1074
Db 1087 TACAGTGCATCAAGCGGTGGAGTAGCACTCAACATATGAATTTATGGTGGGACATAGG 1146
QY 1075 ATTCAATCTCGCCCAATAGGAGCGGATTAATAATCACTCAACGCGATGGTCTACCA---AT 1131
Db 1147 CTTAACTTTCGCGCCCAATAGGAGGACATTAATACTCAACAGAGACTTACTTAATAAT 1206
QY 1132 ACTTCTATTAATCTGTGAAGATTAATCTTCTCTCGAGACGTATATTTGGACTGTAATCA 1191
Db 1207 ACTTCAATTAATCTGTGAACATTAACAGTTTACGTTACGTTCTCGAGACGTTTATAGAACAGATCA 1266
QY 1192 TATGAGAGTGTCTTCTATGGGAAATTTACCTTGAACCTTATCATGTTGCTCCCTACTGTT 1251
Db 1267 AATGAGGACCAATAT-----ACTATTTACTCTCTGTGAATGGAGTACCTTGGCT 1320
QY 1252 AGATTTAATTTTAGGAACCTCTCAGAACTATTTTGAAGAGGTACTGCTAACTATAGTCAA 1311
Db 1321 AGATTTAATTTTAAACCCCTCAGAAATATTTATGAAGAGGCGCACTACCTACAGTCAA 1380
QY 1312 CCCTATCAGTCACTGGGCTTCAATTTAAAGATTTACAGAACTGAAATACCAACAGAAACA 1371
Db 1381 CCGTATCAGGGAGTTGGGATTCANTTATTTGATTCAGAACTGAAATTTACCAACAGAAACA 1440
QY 1372 ACAGAACGACCAATTTATGAATCATATAGTCAATAGTATCTCAGATAGGCTCATTTTCA 1431
Db 1441 ACAGAACGACCAATTTATGAATCATATAGTCAATAGTATCTCAGATAGGCTCATATCA 1500
QY 1432 CAATCTAGGCGTCACTGACAGTATATCTTGGAGCGCCCGTAGTGCAGATCGTACAAAT 1491
Db 1501 GGAACACACTTTGAGAGCACCGTCTATTTCTGGAGCGATCGTAGTGCAGATCGTACGAAT 1560
QY 1492 ACCATTTAGTTTCAGATGACATACAAATACCATTTGGTAAAGCACTGAACTTCAATTTCA 1551
Db 1561 ACATTTGAGACCAATTAGAATTAACAAATACCATTTGGTAAAGCACTGAACTTCAATTTCA 1620
QY 1552 GSTACCTCTGTAGTCAGTGGGCCAGGATTTACAGAGGGGATATTAATCCGAACCTAAAGTT 1611
Db 1621 CGTGTACTTGTGTTGGAGGGCCAGGATTTACAGTGGGGATATCCCTTCGTAGAACAAAT 1680
QY 1612 AATGCTAGTGTACTAAGTATGGGCTTTAATTTTAATATATCATCATTTACAGCGGTATCGC 1671
Db 1681 ACGGGTACATTTGGAGATATACGATTAATAATTAATGTCCTATTATCCCAAGATATCGC 1740
QY 1672 GTGAGAGTTCGTTATGCTGCTTCTCAACAATGCTCTGAGGGTAACTGTGCGGAGGAGT 1731
```

Db 1741 GTAAGGATTCGTATGCTCTACTACAGATTTTCAATTTTTCACGAGAAATTAATGGAACC 1800
Qy 1732 ACTACTTTTGGATCAAGGATTCCTAGTACTATGAGTGCAAAATGAGTCTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATTTGGTAATTTCTCAAGAACTATGAATAGGGGGGTAATTTAGAAATATAGA 1860
Qy 1792 TCATTTAGATTTGCGAAATTTCTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTCTGGA 1851
Db 1861 AGTTTTAGAACTGCGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCCCAAGCAATTC 1920
Qy 1852 ATAAGTATAAGTAATAATCGAGGTAGACAAAGCTTTCACTTTGATAAAATTTGAATTCAT 1911
Db 1921 ACAATTTGGGTGCTCAGAGTTTTTCAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1980
Qy 1912 CCAATTTACTGCAACCTTCGAAGCAGAAATACGATTTTAGAAAGGGCGCAGAGGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAAATTTGAGCGAGATATGATTTAGAAAGAGCAAAAGGCGGTGAAT 2040
Qy 1972 GCTCTGTTTACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTTATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTGAAACAGATGTGACAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTAGATGAAGAAAGAGA 2091
Db 2101 GACCAAGTGTCCAAATATGTTGGCATGTTTATCAGATGAATTTTGTCTGATGAGAAGCGA 2160
Qy 2092 GAATTTACTTTGAGAAAGTGAATATCGAAACGACTCAGTGATGAAGAAACCTTACTCCAA 2151
Db 2161 GAATTTATTTGAGAAAGTGAATATCGAAAGCGACTCAGTGATGAAGAAACCTTACTCCAA 2220
Qy 2152 GATCCAAACTTCAATCCATCAATGAAGCAACAGACCTTCAATCTACTAATGAGCAATCG 2211
Db 2221 GATCCAAACTTCAATCCATCAAGTGGCAATTTAAGTTTCGCATCCATCGATGGAACAATCA 2280
Qy 2212 AATTTCAATCTATCCATGAACAACTCTGAAATGATGTTGGGGAAGTGAGAAATATACA 2271
Db 2281 AACTTTCCCTCTATTAATGAGCTATCTGAACATGATGTTGGGGAAGTGCGAATGTTACC 2340
Qy 2272 ATCCAGGAAGGAATGAGCTATTTAAAGAGAAATACGTCACACTACCGGGGACTTTTAAT 2331
Db 2341 ATTCAAGGAAGGAATGAGCTATTTAAAGAAATTTAGCAACTACCGGGGACTTTTAAT 2400
Qy 2332 GAGTGTATCCGAGTATTTATATCAAAATAATGAGAGTGCAGAAATTTAAAGCTTTATACT 2391
Db 2401 GAGTGTATCCAAATTTATATCAAAATAATGAGAGTGCAGAAATTTAAAGCTTTATAGC 2460
Qy 2392 CGCTACCAATTTAAGAGGTATTTGAAGATGATCAAGATTTAGAGATATATTTGATTCGT 2451
Db 2461 CGCTATCAATTTAAGAGGTATTTGAAGATGATCAAGATTTAGAGATTTATTTAATTCGT 2520
Qy 2452 TATAATGCGAAACATGAAACATTTGATGTTCCAGGTACCGATCGGTATGCGCGCTTCA 2511
Db 2521 TACAATGCAAGCAATGAAACATTTGATGTTCCAGGTACCGATTTCCCTATGCGCGCTTCA 2580
Qy 2512 GTTGAAGCCCAATCCGAGGTGCGGAGAACCGAATCGATCGGACCACTTTTGAATGG 2571
Db 2581 GTTGAAGCCCAATCCGAGGTGCGGAGAACCGAATCGATCGGACCACTTTTGAATGG 2640
Qy 2572 AATCTGATCTAGATTTGTTCTGCAAGATGAGAAATAATGTGCGCATCATTTCCCATCAT 2631
Db 2641 AATCTGATCTAGATTTGTTCTGCAAGATGAGAAATAATGTGCGCATCATTTCCCATCAT 2700
Qy 2632 TTCTCTTTGGATTTGATTTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 2691
Db 2701 TTCTCTTTGGATTTGATTTGATGTTGATGATGATGATGATGATGATGATGATGATGATG 2760
Qy 2692 GTATTTCAAGATTTAAGACGAGGAAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTGAA 2751
Db 2761 GTATTTCAAGATTTAAGACGAGGAAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTGAA 2820
Qy 2752 GAGAAACCAATTTATAGGAAGACATGCTGCTCGTGTGAAGAGAGAGAGAGAGAGAGAGAG 2811

Db 2821 GAGAAACCAATTTATTTGGAGAGACATGCTGCTGTGTGAAGAGAGCGGAAAAAATTTGGAGA 2880
Qy 2812 GACAAACGTTGAAAACTCAATTTGAAAAACAAACGAGTATATACAGAGCGCAAAAGAGCT 2871
Db 2881 GACAAACGTTGAAAACTCAATTTGAAAAACAAACGAGTATATACAGAGCGCAAAAGAGCT 2940
Qy 2872 GTGGATGCTTTATTTGTAGATTTCTCAATATATATAGATTTCAAGCGGATACAAACATTTGC 2931
Db 2941 GTGGATGCTTTATTTGTAGATTTCTCAATATATATAGATTTCAAGCGGATACAAACATTTGC 3000
Qy 2932 ATGATTCATGCGGCGAGATAAATCTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAAT 2991
Db 3001 ATGATTCATGCGGCGAGATAAATCTGTTTCATCGAATTCGAGAGGCTTATCTTTTCAGAAAT 3060
Qy 2992 TCTGTTATCCCGGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAAGTTCGCAATTTCACT 3051
Db 3061 CTTGTTATCCAGAGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAAGTTCGCAATTTCACT 3120
Qy 3052 GCAATCTCCCTATACGATGCGAGAAATCTGTTTAAAAATGTTGATTTTAAATATGATTTA 3111
Db 3121 GCAATGCTCTTATACGATGCGAGAAATCTGTTTAAAAATGTTGATTTTAAATATGATTTA 3180
Qy 3112 GCATGCTGGAATGTAAAGGGCATGTAGATGTACAAAGAGCCATCACGTTCTGTCTCT 3171
Db 3181 ACATGTTGGAATGTAAAGGGCATGTAGATGTACAAAGAGCCATCATCGTTCTGACCTT 3240
Qy 3172 GTTATCCAGAAATGGGAAGCAGAGTGTCAAAAGCAGTTCCGCTCTGTCCGGGGCGTGC 3231
Db 3241 GTTATCCAGAAATGGGAAGCAGAGTGTCAAAAGCAGTTCCGCTCTGTCCGGGGCGTGC 3300
Qy 3232 TATATCTCTCTGTGTACAGCGGTACAAAGAGGGATATGGAGAGGTTGTGTAAACGATCCAT 3291
Db 3301 TATATCTCTCTGTGTACAGCGGTACAAAGAGGGATATGGAGAGGTTGTGTAAACGATCCAT 3360
Qy 3292 GAATTCGAGAAATACAGAGCAATTAATTAATTTTAAAACTGTGAAGAGAGAGAGTGTAT 3351
Db 3361 GAATTCGAGAAATACAGAGCAATTAATTAATTTTAAAACTGTGAAGAGAGAGAGTGTAT 3420
Qy 3352 CCAACGGATACAGAAACGTTGATTAATGATTTACTGCACCAAGGTACAGCAGTATGTAAT 3411
Db 3421 CCACGGATACAGAAACGTTGATTAATGATTTACTGCACCAAGGTACAGCAGTATGTAAT 3480
Qy 3412 TCCGTAATGCTGGATATGAGGATGCAATATGAAGTTGATCTACAGCATCTGTTAATTAC 3471
Db 3481 TCCGTAATGCTGGATATGAGGATGCAATATGAAGTTGATCTACAGCATCTGTTAATTAC 3540
Qy 3472 AAACCGACTTATGAAGAAACGTTATACAGATGTACGAGAGATAATCATTTGGAATAT 3531
Db 3541 AAACCGACTTATGAAGAAACGTTATACAGATGTACGAGAGATAATCATTTGGAATAT 3600
Qy 3532 GACAGAGGTATGTAATTTATCCACTACCGCTGTTATATGACAAAAGAAATTTAGAA 3591
Db 3601 GACAGAGGTATGTAATTTATCCAGTACCGCTGTTATGTCACAAAAGAAATTTAGAA 3660
Qy 3592 TACTTCCAGAAACCGGATTAAGTATGAGATTTGAGAAACGGAAGGGAAGTTTATT 3651
Db 3661 TACTTCCAGAAACCGGATTAAGTATGAGATTTGAGAAACGGAAGGGAAGTTTATT 3720
Qy 3652 GTAGACAGCTGGAATTTACTCTTTATGGAGAAATAG 3687
Db 3721 GTAGATAGCTGGAACTACTCTCTATGGAAGAAATAG 3756

Search completed: November 21, 2004, 13:33:28
Job time : 1492 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 08:48:00 ; Search time 1565 Seconds
(without alignments)
12367.166 Million cell updates/sec

Title: US-10-614-524-1
Perfect score: 3687
Sequence: 1 ttgacttcaatagaaaaa.....tactccttatggaggaatag 3687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1:	Geneseqn1980s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3687	100.0	3687	4	Aah28240 Nucleotid
2	3666.2	99.4	4173	4	Aas02482 B. thurin
3	3266.6	88.6	3684	2	Aav15222 Bacillus
4	3266.6	88.6	3684	4	Aaa89221 Bacillus
5	3260.2	88.4	3684	4	Aas02489 B. thurin
6	3132	84.9	4074	2	Aac56804 Bacillus
7	2687.2	72.9	3932	2	Aag64112 cryET5 ge
8	2687.2	72.9	3932	2	Aaz09160 B. thurin
9	2687.2	72.9	3934	2	Aat95051 DNA encod
10	2687.2	72.9	3934	2	Aat68434 CryET5 ge
11	2687.2	72.9	3934	2	Aav83927 DNA encod
12	2684.2	72.8	3687	12	Adk98478 B thuring
13	2630.2	71.3	3627	4	Aas04854 Bacillus
14	2544.4	69.0	3663	4	Aas00422 B. thurin
15	2541.2	68.9	3663	4	Aas00421 B. thurin
16	2384.2	64.7	3624	4	Aas04853 Bacillus
17	2004.8	54.4	3651	4	Aas04852 Bacillus
18	1793.4	48.6	3522	2	Aat16558 B. t. toxi
19	1791.8	48.6	3522	2	Aaq06830 Sequence
20	1791.8	48.6	3522	2	Aat49592 Bacillus
21	1790.2	48.6	3522	2	Aaq26928 DNA encod

22	1662	45.1	4106	2	AAV16214	Aav16214 DNA encod
23	1662	45.1	4106	2	AAV09318	Aav09318 Bacillus
24	1581.6	42.9	3558	2	AAx60619	Aax60619 Polynucle
25	1501	40.7	3766	2	AAQ77858	AAQ77858 Crystal p
26	1438.6	39.0	8854	2	AAZ20086	AAZ20086 DNA encod
27	1432.4	38.9	3713	2	AAQ64111	AAQ64111 cryET4 ge
28	1432.4	38.9	3713	2	AAAT95050	AAAT95050 DNA encod
29	1432.4	38.9	3713	2	AAAT68435	AAAT68435 CryET4 ge
30	1432.4	38.9	3713	2	AAV83926	AAV83926 DNA encod
31	1432.4	38.9	3713	2	AAZ09159	AAZ09159 B. thurin
32	1411.4	38.3	4020	2	AAQ14648	AAQ14648 cryIF gen
33	1408.4	38.2	3522	2	AAQ10182	AAQ10182 Lepidopte
34	1408.4	38.2	3522	2	AAQ47291	AAQ47291 Delta end
35	1408.4	38.2	3522	2	AAV62081	AAV62081 Plasmid p
36	1408.4	38.2	3522	2	AAx60606	AAx60606 Wild-type
37	1408.4	38.2	3522	4	AAx00565	AAx00565 B. thurin
38	1406.8	38.2	3522	2	AAAT18722	AAAT18722 Codon-rw
39	1405.2	38.1	3522	2	AAAT05270	AAAT05270 CryIF tox
40	1402.2	38.0	5170	12	ADK98488	ADK98488 B thuring
41	1402.2	38.0	5600	12	ADK98490	ADK98490 B thuring
42	1402.2	38.0	6600	12	ADK98482	ADK98482 B thuring
43	1402.2	38.0	7000	12	ADK98485	ADK98485 B thuring
44	1402	38.0	3534	2	AAV31162	AAV31162 Bacillus
45	1402	38.0	3534	6	ABK11833	ABK11833 DNA encod

ALIGNMENTS

RESULT 1
AAH28240
ID AAH28240 standard; DNA; 3687 BP.
XX
AC AAH28240;
XX
DT 05-SEP-2001 (first entry)
XX
DE Nucleotide sequence of a Cry1Bf insecticidal protein.
XX
KW Cry1Bf; insecticidal protein; Cry1Jd; Cry9Pa; lepidoptera; coleoptera;
insect pest; transgenic plant; ss.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
CDS 1..3687
FT /*tag= a
FT /product= "Cry1Bf insecticidal protein"
XX
FN WO200147952-A2.
XX
PD 05-JUL-2001.
XX
PF 19-DEC-2000; 2000WO-EP013184.
XX
PR 28-DEC-1999; 99US-0173387P.
XX
PA (AVET) AVENTIS CROPS SCIENCE NV.
XX
PI Arnaut G, Boets A, Damme N, Mathieu E, Vanneste S, Van Rie J;
XX
DR WPI; 2001-425619/45.
XX
PT Novel insecticidal proteins Cry1Jd, Cry9Pa and Cry1Bf, derived from
Bacillus thuringiensis, useful for controlling insects in plants.
XX
PS P-PSDB; AAB84628.
XX
CC Claim 5; Page 30-36; 65pp; English.
XX
CC The present sequence encodes Cry1Bf, an insecticidal protein derived from
Bacillus thuringiensis. The specification also describes Cry1Jd and
Cry9Pa. The Cry proteins have activity against lepidopteran and
coleopteran insect pests. Cry1Bf, Cry1Jd and Cry9Pa polynucleotides and

CC polypeptides are useful for obtaining a plant with resistance to insects.
CC Cry polynucleotides are useful for producing transgenic plants which are
CC resistant to insects

XX	SQ	Sequence	3687 BP; 1219 A; 640 C; 805 G; 1023 T; 0 U; 0 Other;	
		Query Match	100.0%; Score 3687; DB 4; Length 3687;	
		Best Local Similarity	100.0%; Pred. No. 0;	
		Matches 3687; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	TTGACTTCAAAATAGGAAAAATGAGAAATGAAATTAATAATGCTTTATCGAATTCAGCTGTA	60	
Db	1	TTGACTTCAAAATAGGAAAAATGAGAAATTAATAATGCTTTATCGAATTCAGCTGTA	60	
Qy	61	TCGAATCATTCACACAAATGGATCTATACACAGATGCTCGTATTGAGGATTCCTTTGTGT	120	
Db	61	TCGAATCATTCACACAAATGGATCTATACACAGATGCTCGTATTGAGGATTCCTTTGTGT	120	
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTGTGTTAGCGCATCAACAGTCCAAACGGGTATT	180	
Db	121	ATAGCCGAGGGGAATATATCAATCCACTTGTGTTAGCGCATCAACAGTCCAAACGGGTATT	180	
Qy	181	AACATAGCTGGTAGAATPACTAGTGTATTAGGGGTACCGTTTGTCTGGACAAATAGCTAGT	240	
Db	181	AACATAGCTGGTAGAATPACTAGTGTATTAGGGGTACCGTTTGTCTGGACAAATAGCTAGT	240	
Qy	241	TTTTATAGTTTCTTGTGTGGTAATTATGCCCCCGCGGACAGATCAGTGGGAAATTTTC	300	
Db	241	TTTTATAGTTTCTTGTGTGGTAATTATGCCCCCGCGGACAGATCAGTGGGAAATTTTC	300	
Qy	301	CTAGAAACATGTCGAAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360	
Db	301	CTAGAAACATGTCGAAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360	
Qy	361	CTTGCTCGAATTAAGAATTTAGAGATTCCTTTAGAGCCTATCAACAGTCACCTTGAAGAT	420	
Db	361	CTTGCTCGAATTAAGAATTTAGAGATTCCTTTAGAGCCTATCAACAGTCACCTTGAAGAT	420	
Qy	421	TGGCTAGAAACCGTGATGATCAAGACGAGAGTGTCTTTATACCCAAATATATAGCC	480	
Db	421	TGGCTAGAAACCGTGATGATCAAGACGAGAGTGTCTTTATACCCAAATATATAGCC	480	
Qy	481	TTAGAAACCTTGATTTCTTAATGCGATGCCCTTTTCGCAATTAGAAACCAAGAAGTTCCA	540	
Db	481	TTAGAAACCTTGATTTCTTAATGCGATGCCCTTTTCGCAATTAGAAACCAAGAAGTTCCA	540	
Qy	541	TTATTAAATGATATGCTCAAGCTGCAAAATTTACACTATATATATGAGAGATGCCCTCT	600	
Db	541	TTATTAAATGATATGCTCAAGCTGCAAAATTTACACTATATATATGAGAGATGCCCTCT	600	
Qy	601	CTTTTTGGTAGTCAATTTGGGCTTACATCGCAGGAAATCAACGTTTATTATGAGCGCCAA	660	
Db	601	CTTTTTGGTAGTCAATTTGGGCTTACATCGCAGGAAATCAACGTTTATTATGAGCGCCAA	660	
Qy	661	GTGGAAACAAACGAGAGATTATCCGACTATTGCGTAGAATGGTATTAATACAGGCTTAAAT	720	
Db	661	GTGGAAACAAACGAGAGATTATCCGACTATTGCGTAGAATGGTATTAATACAGGCTTAAAT	720	
Qy	721	AGCTTAGAGGGGACAAATGCCGCAAGTTGGGTGCGTTAATCAATTCGCTAGAGATCTA	780	
Db	721	AGCTTAGAGGGGACAAATGCCGCAAGTTGGGTGCGTTAATCAATTCGCTAGAGATCTA	780	
Qy	781	ACGTTAGGGGTATTAGATCTAGTGGCACATATCCCAAGCTATGACACTCGCACTTATCCA	840	
Db	781	ACGTTAGGGGTATTAGATCTAGTGGCACATATCCCAAGCTATGACACTCGCACTTATCCA	840	
Qy	841	ATAAATAACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTTGGAGCAACAGG	900	
Db	841	ATAAATAACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTTGGAGCAACAGG	900	
Qy	901	GTAATAATGCGAGTATGAATTTGGTATTAATAATGCAACCTTCGTTTCCGCTATAGAG	960	
Db	901	GTAATAATGCGAGTATGAATTTGGTATTAATAATGCAACCTTCGTTTCCGCTATAGAG	960	

Qy	961	ACTGCGGTTATCCGAAGCCCGCATCTACTAGTTGATTTTCTAGAACAACTTACAAATTTTAGC	1020	
Db	961	ACTGCGGTTATCCGAAGCCCGCATCTACTAGTTGATTTTCTAGAACAACTTACAAATTTTAGC	1020	
Qy	1021	ACTTCATACGATGGAGTGTCTACTAGGCATATGACTTACTTGGCGGGGACACAAATTCAA	1080	
Db	1021	ACTTCATACGATGGAGTGTCTACTAGGCATATGACTTACTTGGCGGGGACACAAATTCAA	1080	
Qy	1081	TCTCGGCCAATAGGAGCGGATTAATACCTCAACGCATGGGTCTACCAATACTTCTTATT	1140	
Db	1081	TCTCGGCCAATAGGAGCGGATTAATACCTCAACGCATGGGTCTACCAATACTTCTTATT	1140	
Qy	1141	AATCCTGTAAGATTATCTTCTCGAGAGCTATATGGACTGAATCATATGACGGA	1200	
Db	1141	AATCCTGTAAGATTATCTTCTCGAGAGCTATATGGACTGAATCATATGACGGA	1200	
Qy	1201	GTGCTTCTATGGGAATTTACCTTGAACCTATTCTATGGTGTCCCTACTCTGTAGATTAAAT	1260	
Db	1201	GTGCTTCTATGGGAATTTACCTTGAACCTATTCTATGGTGTCCCTACTCTGTAGATTAAAT	1260	
Qy	1261	TTTAGGAACCCCTCAGAAATCTTTTGAAGAGGTACTGTAACTATATAGTCAACCCCTATGAG	1320	
Db	1261	TTTAGGAACCCCTCAGAAATCTTTTGAAGAGGTACTGTAACTATATAGTCAACCCCTATGAG	1320	
Qy	1321	TCACCTGGGCTTCNAATTAAGATTCAGAACTGAATTAACCAAGAAACAAACAGAACGA	1380	
Db	1321	TCACCTGGGCTTCNAATTAAGATTCAGAACTGAATTAACCAAGAAACAAACAGAACGA	1380	
Qy	1381	CCAAATTAAGAACTCATATAGTCTCATAGGCTCATCTACATAGGCTCAATTTACAACTAGG	1440	
Db	1381	CCAAATTAAGAACTCATATAGTCTCATAGGCTCATCTACATAGGCTCAATTTACAACTAGG	1440	
Qy	1441	GTGCAATGTAACAGTATATCTTTGGAGCGCACCGTAGTGAGATCGTACAAATACCATTAGT	1500	
Db	1441	GTGCAATGTAACAGTATATCTTTGGAGCGCACCGTAGTGAGATCGTACAAATACCATTAGT	1500	
Qy	1501	TCAGATAGCATTAACCAAAATACCAATTCATTTGTAATAATCAATTCAACTTAACTCAGTACT	1560	
Db	1501	TCAGATAGCATTAACCAAAATACCAATTCATTTGTAATAATCAATTCAACTTAACTCAGTACT	1560	
Qy	1561	GTAGTCAGTGGGCCAGGATTTACAGAGGGGATATAATCCGAACCTAACCTTAACTAGT	1620	
Db	1561	GTAGTCAGTGGGCCAGGATTTACAGAGGGGATATAATCCGAACCTAACCTTAACTAGT	1620	
Qy	1621	GTACTAAGTATGGGTCTTAATTTAATAATACATCATTAACAGCGGTATCGCGTGAAGTT	1680	
Db	1621	GTACTAAGTATGGGTCTTAATTTAATAATACATCATTAACAGCGGTATCGCGTGAAGTT	1680	
Qy	1681	CGTTATGCTGCTTCTCAAAACATGGTCTGAGGTAACCTGTCGAGGGAGTACTACTTTT	1740	
Db	1681	CGTTATGCTGCTTCTCAAAACATGGTCTGAGGTAACCTGTCGAGGGAGTACTACTTTT	1740	
Qy	1741	GATCAAGGATCCCTAGTACTATGAGTGCAAAATGAGTCTTTGACATCTCAATCATTTAGA	1800	
Db	1741	GATCAAGGATCCCTAGTACTATGAGTGCAAAATGAGTCTTTGACATCTCAATCATTTAGA	1800	
Qy	1801	TTTGCAGAAATTTCTGTAGGTATTAGTGATCTGGCAGTCAAACTGCTGGAATAAGTATA	1860	
Db	1801	TTTGCAGAAATTTCTGTAGGTATTAGTGATCTGGCAGTCAAACTGCTGGAATAAGTATA	1860	
Qy	1861	AGTAAATATGCAAGTAGACAAAACGTTTCACTTTGATAAAAATGAATTTCAATCCAAATTA	1920	
Db	1861	AGTAAATATGCAAGTAGACAAAACGTTTCACTTTGATAAAAATGAATTTCAATCCAAATTA	1920	
Qy	1921	GCAAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGCGGTGTAATGCTCTGTTT	1980	
Db	1921	GCAAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGCGGTGTAATGCTCTGTTT	1980	
Qy	1981	ACTTAATACCAATCCCAAGAGATTTGAAAACAGATGTGACAGATTTATCATATTGATCAAGTA	2040	
Db	1981	ACTTAATACCAATCCCAAGAGATTTGAAAACAGATGTGACAGATTTATCATATTGATCAAGTA	2040	

2041 TCCAAATTTAGTCGCGTGTATTCGGATGAATTCCTGCTAGATGAAAAAGAGAGAAATTAATCTT 2100
2041 TCCAAATTTAGTCGCGTGTATTCGGATGAATTCCTGCTAGATGAAAAAGAGAGAAATTAATCTT 2100
2101 GAGAAAGTGAATATATCGGAAACGACTCAGTGATGAAGAACTTACTCCAAGATCCAAAC 2160
2101 GAGAAAGTGAATATATCGGAAACGACTCAGTGATGAAGAACTTACTCCAAGATCCAAAC 2160
2161 TCCAAATTCATCAATAAGCAACACGACTTCATATCTACTTAATGAGCAATCGAATTTTACA 2220
2161 TCCAAATTCATCAATAAGCAACACGACTTCATATCTACTTAATGAGCAATCGAATTTTACA 2220
2221 TCTATCCATGAACAACTCTGAAACGACTCAGTGATGAAGAACTTACTCCAAGATCCAAAC 2280
2221 TCTATCCATGAACAACTCTGAAACGACTCAGTGATGAAGAACTTACTCCAAGATCCAAAC 2280
2281 GGAATGACGTTATTTAAAGAGAAATTAACGTCACACTACCGGGGACCTTTTAATGAGTGTAT 2340
2281 GGAATGACGTTATTTAAAGAGAAATTAACGTCACACTACCGGGGACCTTTTAATGAGTGTAT 2340
2341 CCGAGCTATTTATATCAAAAAATAGAGAGTCGGAATTAAGAGCTTATATCTCGCTACCAA 2400
2341 CCGAGCTATTTATATCAAAAAATAGAGAGTCGGAATTAAGAGCTTATATCTCGCTACCAA 2400
2401 TTAAGAGGCTATTTCAAGATAGTCAAGATTTTGAAGATATATTTGATTCGTTTAAATGCG 2460
2401 TTAAGAGGCTATTTGAAGATAGTCAAGATTTTGAAGATATATTTGATTCGTTTAAATGCG 2460
2461 AAACATGAACATTTGATGATTCAGGTACCGTACCGAGTCCGATGATGCGCGCTTCAGTTGAAAGC 2520
2461 AAACATGAACATTTGATGATTCAGGTACCGTACCGAGTCCGATGATGCGCGCTTCAGTTGAAAGC 2520
2521 CCAATCGGAGGTGCGGAAACCGAATCGATCGGACCACTTTTGAATGGAATCCTGAT 2580
2521 CCAATCGGAGGTGCGGAAACCGAATCGATCGGACCACTTTTGAATGGAATCCTGAT 2580
2581 CTAGATTTGTTCTCGAGAGATGGAGAAAATGTCGCAATCAATTCCTCTTTG 2640
2581 CTAGATTTGTTCTCGAGAGATGGAGAAAATGTCGCAATCAATTCCTCTTTG 2640
2641 GATATTTGATTTGGATGACAGACTTCGATGAGATCTAGGCGTGTGGTGTATTCAG 2700
2641 GATATTTGATTTGGATGACAGACTTCGATGAGATCTAGGCGTGTGGTGTATTCAG 2700
2701 ATTAAGACGAGAAAGTTCATCAAGACTTAGGAAATCTGGAATTTATTAAGAGAAACCA 2760
2701 ATTAAGACGAGAAAGTTCATCAAGACTTAGGAAATCTGGAATTTATTAAGAGAAACCA 2760
2761 TTATTAGGAGAGCACTGTCTCGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
2761 TTATTAGGAGAGCACTGTCTCGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
2821 GAAAACTTAAATTTGAAACAAACAGATATATACAGAGGCAAAAGAGAGCTGTGATGCT 2880
2821 GAAAACTTAAATTTGAAACAAACAGATATATACAGAGGCAAAAGAGAGCTGTGATGCT 2880
2881 TTATTGATGATTTCTCAATATATAGATTTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
2881 TTATTGATGATTTCTCAATATATAGATTTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
2941 GCGGAGATAAATCTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTCTGTTATC 3000
2941 GCGGAGATAAATCTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTCTGTTATC 3000
3001 CCGGCTGTAATTCGGAATTTTGAAGATTTAGAAAGTTCGATTTATCACTGCAATCTCC 3060
3001 CCGGCTGTAATTCGGAATTTTGAAGATTTAGAAAGTTCGATTTATCACTGCAATCTCC 3060
3061 CTATACGATCGGAGAAATTCGTTTAAATGATGATTTTAAATGGAATTTAGCATGCTGG 3120
3061 CTATACGATCGGAGAAATTCGTTTAAATGATGATTTTAAATGGAATTTAGCATGCTGG 3120
3121 AATGTAAGAGGCGATGTAGATGTACACAGAGCCATCCCGTTCCTGCTGTTATCCCA 3180

Db 3121 AATGTAAGAGGCGATGTAGATGTACACAGAGCCATCCCGTTCCTGCTGTTATCCCA 3180
Qy 3181 GAATGGAGAGCAGAAAGTGTCAAGACAGTTCGCTGTCTCGGGGCGTGGCTATATCTTC 3240
Db 3181 GAATGGAGAGCAGAAAGTGTCAAGACAGTTCGCTGTCTCGGGGCGTGGCTATATCTTC 3240
Qy 3241 CGTGTCAAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAAATCGAG 3300
Db 3241 CGTGTCAAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAAATCGAG 3300
Qy 3301 AACAAATACAGACGAACTAAAAATTTAAAACTGTGAAGAGAGAAAGTGTATCCAAACGAT 3360
Db 3301 AACAAATACAGACGAACTAAAAATTTAAAACTGTGAAGAGAGAAAGTGTATCCAAACGAT 3360
Qy 3361 ACAGGACGCTGAATGATTTACTGCAACCAAGGTACAGCAGTATGTAATTTCCCGTAA 3420
Db 3361 ACAGGACGCTGAATGATTTACTGCAACCAAGGTACAGCAGTATGTAATTTCCCGTAA 3420
Qy 3421 GCTGATATGAGGATGCATATGAAAGTTCATCTACAGCATCTGTTAAATTTACAAACCGACT 3480
Db 3421 GCTGATATGAGGATGCATATGAAAGTTCATCTACAGCATCTGTTAAATTTACAAACCGACT 3480
Qy 3481 TATGAAGAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTAATATGACAGAGGG 3540
Db 3481 TATGAAGAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTAATATGACAGAGGG 3540
Qy 3541 TATGTAATTTATCCCACTACCGACTGCTGTTATATGACAAAAGATTTAGAAATTTCCCA 3600
Db 3541 TATGTAATTTATCCCACTACCGACTGCTGTTATATGACAAAAGATTTAGAAATTTCCCA 3600
Qy 3601 GAAACCGATAAGGTATGATGATTCGAAATTCGAGAAACGGAAGGAAAGTATTGTTAGACAGC 3660
Db 3601 GAAACCGATAAGGTATGATGATTCGAAATTCGAGAAACGGAAGGAAAGTATTGTTAGACAGC 3660
Qy 3661 GTGGAATTTACTCTTTATGAGAGAAATAG 3687
Db 3661 GTGGAATTTACTCTTTATGAGAGAAATAG 3687

RESULT 2
AAS02482
ID AAS02482 standard; DNA; 4173 BP.
XX
AC AAS02482;
XX
DT 29-AUG-2001 (first entry)
XX
DE
XX
KW Delta endotoxin; Lepidopteran-active; crystal protein; insecticide;
KW transgenic plant; corn; wheat; soybean; oat; cotton; rice; yre; sorghum;
KW sugarcane; tomato; tobacco; kapok; flax; potato; barley; turf grass;
KW pasture grass; berry; fruit; legume; vegetable; ornamental plant; shrub;
KW cactus; tree cell; gypsy moth; looper; tobacco budworm; spruce budworm;
KW cotton leaf perforator; CryE64; ds.
XX
OS Bacillus thuringiensis.
XX
Key Location/Qualifiers
FT CDS 1..3471
FT /tag= a
FT /product= "CryE64"
FT /trans_except= (pos:1..3,aa:Met)
XX
WO200119859-A2.
PD 22-MAR-2001.
XX
PF 13-SEP-2000; 2000WO-US025361.
XX
PR 15-SEP-1999; 99US-0153995P.
XX

PA	(MONS) MONSANTO CO.	
XX		
PI	Baum JA, Chu C, Donovan WP, Gilmer AJ, Rupar MJ;	
XX		
DR	WPI; 2001-281518/29.	
XX	P-PSDB; AAU02039.	
XX		
PT	Lepidopteran-active Bacillus thuringiensis delta-endotoxin polypeptides	
PT	and the polynucleotides that encode them, useful for increasing the	
XX	insect resistance of plant.	
XX		
PS	Claim 17; Page 133-138; 173pp; English.	
XX		
CC	The sequence encodes a B. thuringiensis Lepidopteran-active delta-	
CC	endotoxin, crystal protein CryIb4. The Lepidopteran-active B.	
CC	thuringiensis delta-endotoxin polypeptides may be used as compositions	
CC	that are applied to plant crops to protect them from insect damage. The	
CC	polynucleotides may be used in the production of transgenic plants that	
CC	express the insecticidal polypeptides and consequently have improved	
CC	insect resistance compared to non-transformed plants. Monocotyledonous or	
CC	dicotyledonous plants may be protected in this way, for example corn,	
CC	wheat, soybean, oat, cotton, rice, rye, sorghum, sugarcane, tomato,	
CC	tobacco, kapok, flax, potato, barley, turf grass, pasture grass, berry,	
CC	fruit, legume, vegetable, ornamental plant, shrub, cactus and/or tree	
CC	cell. A wide range of insects (e.g. gypsy moth, looper, tobacco budworm,	
CC	cotton leaf perforator and spruce budworm) may be affected by application	
CC	of the insecticidal polypeptides (full details given in specification)	
XX		
SQ	Sequence 4173 BP; 1417 A; 705 C; 878 G; 1173 T; 0 U; 0 Other;	
	Query Match 99.4%; Score 3666.2; DB 4; Length 4173;	
	Best Local Similarity 99.6%; Pred. No. 0;	
	Matches 3674; Conservative 0; Mismatches 13; Indels 0; Gaps 0;	
Qy	1 TTGACTTCAAAATAGGAAAAATGAGAAATTAATAATGCTTTATCGATTCACGCTGTA 60	
Db	1 TTGACTTCAAAATAGGAAAAATGAGAAATTAATAATGCTTTATCGATTCACGCTGTA 60	
Qy	61 TCGAATCATTTCCACAAATGGATCTATACACGATGCTCGTATTCAGGATCTTTGTGT 120	
Db	61 TCGAATCATTTCCACAAATGGATCTATACACGATGCTCGTATTCAGGATCTTTGTGT 120	
Qy	121 ATAGCCGAGGGGAATATATCAATCCACTTGTAGGCGATCAACAGTCCAAAGCGGTAT 180	
Db	121 ATAGCCGAGGGGAATATATCAATCCACTTGTAGGCGATCAACAGTCCAAAGCGGTAT 180	
Qy	181 AACATAGCTGGTAGAATACATAGGTGTATAGGGGTACCGTTTCTGGACAAATAGTAGT 240	
Db	181 AACATAGCTGGTAGAATACATAGGTGTATAGGGGTACCGTTTCTGGACAAATAGTAGT 240	
Qy	241 TTTTATATGTTTCTTGTGTGTAATATATGCCCCCGGCGAGAGATCAGTGGGAAATTTTC 300	
Db	241 TTTTATATGTTTCTTGTGTGTAATATATGCCCCCGGCGAGAGATCAGTGGGAAATTTTC 300	
Qy	301 CTAGAACATGCGAACAACTTATAATCAACAAATACAGAAATGCTAGGAATACGGCA 360	
Db	301 CTAGAACATGCGAACAACTTATAATCAACAAATACAGAAATGCTAGGAATACGGCA 360	
Qy	361 CTTGCTCGATTACAGGTTTAGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420	
Db	361 CTTGCTCGATTACAGGTTTAGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420	
Qy	421 TGGCTAGAAAACCGTGATGATCAAGAACGAGAAAGTGTCTTTATACCCAAATATATAGCC 480	
Db	421 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAAGTGTCTTTATACCCAAATATATAGCC 480	
Qy	481 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAAGTTCCA 540	
Db	481 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAAGTTCCA 540	
Qy	541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCTCT 600	
Db	541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCTCT 600	

Qy	601 CTTTTTGGTAGTAATTTGGGCTTACATCGCAGGAAATTCACAGTTATTATATGAGCGCAA 660
Db	601 CTTTTTGGTAGTAATTTGGGCTTACATCGCAGGAAATTCACAGTTATTATGAGCGCAA 660
Qy	661 GTGGAAACAACGAGAGATTTATCCGACTATTTGGGTAGATGGTATATACAGGCTCTAAAT 720
Db	661 GTGGAAACAACGAGAGATTTATCCGACTATTTGGGTAGATGGTATATACAGGCTCTAAAT 720
Qy	721 AGCTTCAGAGAGGACAAATGCCCAAGTTCGGGTGCGTTATAATCAATTCGCTAGAGATCTA 780
Db	721 AGCTTCAGAGAGGACAAATGCCCAAGTTCGGGTGCGTTATAATCAATTCGCTAGAGATCTA 780
Qy	781 ACCTTAGGGGTATATAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 840
Db	781 ACCTTAGGGGTATATAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 840
Qy	841 ATAAATACAGTGCCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTCGAGCAACAGGG 900
Db	841 ATAAATACAGTGCCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTCGAGCAACAGGG 900
Qy	901 GTAAATATGCGCAAGTATGATGATGTCATATTAATGCAACCTTCGTTTCCGCTATAGAG 960
Db	901 GTAAATATGCGCAAGTATGATGTCATATTAATGCAACCTTCGTTTCCGCTATAGAG 960
Qy	961 ACTGGGGTTATCCGAAGCCCGCATCTACTTTGATTTTCTAGAACAACTTACAAATTTTAGC 1020
Db	961 ACTGGGGTTATCCGAAGCCCGCATCTACTTTGATTTTCTAGAACAACTTACAAATTTTAGC 1020
Qy	1021 ACTTCATACAGATGAGTGTCTACTAGGCATATAGCTTACTGGGGGGGACACAAATTTCAA 1080
Db	1021 ACTTCATACAGATGAGTGTCTACTAGGCATATAGCTTACTGGGGGGGACACAAATTTCAA 1080
Qy	1081 TCTCGGCCCAATAGGAGGGGATTAATATCCTCAAGCGATGGGTCTACCAATATCTTCTATT 1140
Db	1081 TCTCGGCCCAATAGGAGGGGATTAATATCCTCAAGCGATGGGTCTACCAATATCTTCTATT 1140
Qy	1141 AATCCTGTAGATTAATCATCTCTCTCGAGACGTATATTGGACGTGAATCATATGACGGA 1200
Db	1141 AATCCTGTAGATTAATCATCTCTCTCGAGACGTATATTGGACGTGAATCATATGACGGA 1200
Qy	1201 GTGCTTCTATGGGAAATTTACCTTCAACCTATTCATGGTGTCCCTACTCTGTAGATTTAAT 1260
Db	1201 GTGCTTCTATGGGAAATTTACCTTCAACCTATTCATGGTGTCCCTACTCTGTAGATTTAAT 1260
Qy	1261 TTTAGGAACCTTCAGAAATCTTTTGAAGAGGTACTGTCTAATAGTCAACCCCTATGAG 1320
Db	1261 TTTAGGAACCTTCAGAAATCTTTTGAAGAGGTACTGTCTAATAGTCAACCCCTATGAG 1320
Qy	1321 TCACCTGGGCTTCAATTTAAAGATTCAGAACTGAATTTACCAACGAGAAACAAAGAACGA 1380
Db	1321 TCACCTGGGCTTCAATTTAAAGATTCAGAACTGAATTTACCAACGAGAAACAAAGAACGA 1380
Qy	1381 CCAAAATTTAGATCATATAGTCAATAGTGTATCTCAATAGGGCTCATTTTCAAACTCTAGG 1440
Db	1381 CCAAAATTTAGATCATATAGTCAATAGTGTATCTCAATAGGGCTCATTTTCAAACTCTAGG 1440
Qy	1441 GTGCAATGATACAGTATATTTCTTGGACGCAACCGTATGCGAGATCGTACAAATACCAAT 1500
Db	1441 GTGCAATGATACAGTATATTTCTTGGACGCAACCGTATGCGAGATCGTACAAATACCAAT 1500
Qy	1501 TCAGATAGCATTAACAAATACCAATTTGTTAAATCATTTCAACCTTAATTCAGGTACCTCT 1560
Db	1501 TCAGATAGCATTAACAAATACCAATTTGTTAAATCATTTCAACCTTAATTCAGGTACCTCT 1560
Qy	1561 GTAGTCAGTGGCCAGGATTTACAGGAGGGGATTAATCCGAACCTTAAGTAAATGGTAGT 1620
Db	1561 GTAGTCAGTGGCCAGGATTTACAGGAGGGGATTAATCCGAACCTTAAGTAAATGGTAGT 1620
Qy	1621 GTACTAAGTATGGGCTTTAAATTTTAAATATACATCATTTACAGCGGTATCGGTGAGAGTT 1680
Db	1621 GTACTAAGTATGGGCTTTAAATTTTAAATATACATCATTTACAGCGGTATCGGTGAGAGTT 1680

Qy	1681	CgTtATGCTGCTTCtCAAAcAAtGGGTCTGAGGGTAACtGTfCGAGGGAgTACTACTTTT	1740
Db	1681	CgTtATGCTGCTTCtCAAAcAAtGGGTCTGAGGGTAACtGTfCGAGGGAgTACTACTTTT	1740
Qy	1741	GATCAAGGATTCCCTAGTACTATGAGTGCCAATCGATCTTTTGACATCTCAATCATTTAGA	1800
Db	1741	GATCAAGGATTCCCTAGTACTATGAGTGCCAATCGATCTTTTGACATCTCAATCATTTAGA	1800
Qy	1801	TTTTGCAGAAATTTCCCTGTAGGTATTAGTGCATCTGCCAGTCAAACtGTGCTGGAATAAGTATA	1860
Db	1801	TTTTGCAGAAATTTCCCTGTAGGTATTAGTGCATCTGCCAGTCAAACtGTGCTGGAATAAGTATA	1860
Qy	1861	AGTAAATATGCAGGTAGACAACCGTTTCATCTTTGATAAAAATTGAATTCATTTCCAAATTACT	1920
Db	1861	AGTAAATATGCAGGTAGACAACCGTTTCATCTTTGATAAAAATTGAATTCATTTCCAAATTACT	1920
Qy	1921	GCAACCTTCGAAGCAGAGAATACGATTTAGAAAAGCGCGCAAGAGCGGTGAATGCTCTGTTTT	1980
Db	1921	GCAACCTTCGAAGCAGAGAATACGATTTAGAAAAGCGCGCAAGAGCGGTGAATGCTCTGTTTT	1980
Qy	1981	ACTAATACGAATCCAAGAAAGATTGCAAAACAGATGTGCACAGATTATCATATTGATCAAGTA	2040
Db	1981	ACTAATACGAATCCAAGAAAGATTGCAAAACAGATGTGCACAGATTATCATATTGATCAAGTA	2040
Qy	2041	TCCAAATTTAGTGGCGTGTTTATCGSAGTAATTTCTGCTTAGATGAAAAAGAGAGAATTACTT	2100
Db	2041	TCCAAATTTAGTGGCGTGTTTATCGSAGTAATTTCTGCTTAGATGAAAAAGAGAGAATTACTT	2100
Qy	2101	GAGAAAGTGAATATGCCAAACGACTCAGTGATCAAGAGAACCTTACTCCAGATCCCAAC	2160
Db	2101	GAGAAAGTGAATATGCCAAACGACTCAGTGATCAAGAGAACCTTACTCCAGATCCCAAC	2160
Qy	2161	TTCCATCCATCAATTAAGCAACACAGACTTCATATCTACTAATGAGCAATCGAATTTCCACA	2220
Db	2161	TTCCATCCATCAATTAAGCAACACAGACTTCATATCTACTAATGAGCAATCGAATTTCCACA	2220
Qy	2221	TCTATCCATGAACAACTCGAATCGGATGGTGGGAAAGTGAGAACTTACAATCCAGGAA	2280
Db	2221	TCTATCCATGAACAACTCGAATCGGATGGTGGGAAAGTGAGAACTTACAATCCAGGAA	2280
Qy	2281	GGAAATGACGTATTTAAGAGAAATACGTACACACTACCGGGGACTTTTAATGAGTGTTAT	2340
Db	2281	GGAAATGACGTATTTAAGAGAAATACGTACACACTACCGGGGACTTTTAATGAGTGTTAT	2340
Qy	2341	CCGACGTATTTATCAAAAAATAGGAGAGTCCGAAATTAAAGCTTATATCTCGCTACCAA	2400
Db	2341	CCGACGTATTTATCAAAAAATAGGAGAGTCCGAAATTAAAGCTTATATCTCGCTACCAA	2400
Qy	2401	TTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAATCGC	2460
Db	2401	TTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAATCGC	2460
Qy	2461	AAACATGNAACATTGATGTTTCCAGGTACCAGGTCCGTTATGGCGCTTTTCAGTTGAAAGC	2520
Db	2461	AAACATGNAACATTGATGTTTCCAGGTACCAGGTCCGTTATGGCGCTTTTCAGTTGAAAGC	2520
Qy	2521	CCAAATCCGAAGGTGCGGAGAACCAATCGATGCGCACACATTTTGAATGGAATCCCTGAT	2580
Db	2521	CCAAATCCGAAGGTGCGGAGAACCAATCGATGCGCACACATTTTGAATGGAATCCCTGAT	2580
Qy	2581	CTAGATTGTTCTCTCAGAGATGGAGAAAAATGTGCGCATCATTTCCCACATCTTCTTTTG	2640
Db	2581	CTAGATTGTTCTCTCAGAGATGGAGAAAAATGTGCGCATCATTTCCCACATCTTCTTTTG	2640
Qy	2641	GATATTGATATTGATGCAAGACTTTCGATGAGAACTTAGCGGTGTGGGTGGTATTCCAAG	2700
Db	2641	GATATTGATATTGATGCAAGACTTTCGATGAGAACTTAGCGGTGTGGGTGGTATTCCAAG	2700
Qy	2701	ATTAAGACGCGAGGAAGGTCTATGCAAGACTAGGGAATCTGGAAATTTATTGAGAGGAACCA	2760
Db	2701	ATTAAGACGCGAGGAAGGTCTATGCAAGACTAGGGAATCTGGAAATTTATTGAGAGGAACCA	2760
Qy	2761	TTATTAGGAAGACCACTGCTCTGCTGTGAAGAGACAGAGAAAAATTCGAGAGACAACCGT	2820

[illegible]

RESIT.T 3

AAV15222

AAV15222
ID AAV15222 standard: DNA: 3684 BP.

2000

AC AAV15222;

XXXXXX

DT 27-MAY-1998 (first entry)

XX DE Bacillus thuringiensis 158C2c toxin encoding DNA.

DE *Bacillus thuringiensis* 158C2c toxin encoding DNA.

XX Bacillus thuringiensis; 158C2c; toxin; lepidopteran-active; insect;
 KW microbe; transgenic plant; resistant; ss.
 XX Bacillus thuringiensis.
 OS
 XX Key Location/Qualifiers
 XX CDS 1..3684
 FT /*tag= a
 FT /product= "158C2c toxin"
 FT
 XX US5723758-A.
 XX
 XX 03-MAR-1998.
 XX
 XX 23-MAY-1995; 95US-00448170.
 XX
 XX 13-SEP-1991; 91US-00759247.
 XX 01-JUN-1993; 93US-00069902.
 XX (MYCO) MYCOGEN CORP.
 XX
 XX Steelman S, Narva KB, Cummings DA, Payne J, Cannon RJ;
 XX WPI; 1998-178600/16.
 XX P-PSDB; AAW44321.
 XX
 XX DNA encoding Bacillus thuringiensis toxin proteins - for producing
 PT transgenic plants resistant to attack by lepidopteran pests.
 XX
 XX Claim 7; Col 29-34; 25pp; English.
 XX
 XX The present sequence encodes a 158C2c toxin from Bacillus thuringiensis.
 CC The present invention describes isolated nucleic acid sequences encoding
 CC toxins active against lepidopteran insects. The present invention also
 CC describes a recombinant host transformed with a nucleic acid sequence
 CC encoding a toxin. The nucleic acid sequence can be amplified. The product
 CC is used for producing transgenic plants resistant to attack by
 CC lepidopteran pests. N.B. The numbers in the sequence listing do not
 CC correspond to the SEQ ID NO.'s mentioned in the specification e.g. in the
 CC specification SEQ ID NO:3, 4, 5 and 6 are said to correspond to the
 CC nucleic acid sequences encoding the toxins, and SEQ ID NO:7, 8, 9 and 10
 CC are said to correspond to the amino acid sequences of the toxins, whereas
 CC in the sequence listing the amino acid sequence is given after the
 CC nucleic acid sequence encoding it (i.e. SEQ ID NO:3 encodes SEQ ID NO:4)
 XX
 XX Sequence 3684 BP; 1232 A; 628 C; 803 G; 1021 T; 0 U; 0 Other;
 SQ
 Query Match 88.6%; Score 3266.6; DB 2; Length 3684;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 3449; Conservative 0; Mismatches 229; Indels 15; Gaps 3;
 QY 1 TTGACCTCAATAGGAAATGAGATGAATTAATTAATGCTTTATCGATTCCAGCTGTA 60
 DB 1 TTGACCTCAATAGGAAATGAGATGAATTAATTAATGCTTTATCGATTCCAGCTGTA 60
 QY 61 TCGAATCATTCACACAAATGGATCTATCACAGATGCTCGTATTGAGGATTCTTTGTGT 120
 DB 61 TCGAATCATTCGACACAAATGAATCTATCAACGATGCTCGTATTGAGGATGTTGTGT 120
 QY 121 ATAGCCGAGGGGAATTAATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
 DB 121 ATAGCCGAGGGGAATTAATCGATCCATTTGTTAGGCGATCAACAGTCCAAACGGGTATT 180
 QY 181 AACATAGCTGTAGATCTAGTGTATTAGGCGTACCGTTTCTCGACAAATAGTAGT 240
 DB 181 AACATAGCTGTAGATCTAGTGTATTAGGCGTACCGTTTCTCGACAAATAGTAGT 240
 QY 241 TTTTATAGTTTCTTGTGTGAATTAATGCGCCCGGCGGAGAGATCAGTGGGAAATTTTC 300
 DB 241 TTTTATAGTTTCTTGTGTGAATTAATGCGCCCGGCGGAGAGATCCTTGGGAAATTTTC 300
 QY 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATTAACGAAATGCTAGGAATACGGCA 360

DB 301 CTAGAACATGTCGAACAACTTATAAGACAAACAGTAACAGAAATACTAGGGATACGGCT 360
 QY 361 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCCTATCAACAGTCACCTTTGAAGAT 420
 DB 361 CTTGCTCGATTACAAGGTTTAGGAAATTCCTTTAGAGCCCTATCAACAGTCACCTTTGAAGAT 420
 QY 421 TGGCTAGAAAAACCGTGATGATGCAAGAACGAGAGAGTGTCTTTATACCAATATATAGCC 480
 DB 421 TGGCTAGAAAAACCGTGATGATGCAAGAACGAGAGAGTGTCTTTATACCAATATATAGCC 480
 QY 481 TTAGAACTTGTATTTCTTAATCGATGCGCTTTTCGCAATTAGAACAACGAAGTTCCTCA 540
 DB 481 TTAGAACTTGTATTTCTTAATCGATGCGCTTTTCGCAATTAGAACAACGAAGTTCCTCA 540
 QY 541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACTTATTTATTTAGAGAGATGCTCTCT 600
 DB 541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACTTATTTATTTAGAGAGATGCTCTCT 600
 QY 601 CTTTGTGTAGTGAATTTGGGCTTACATCCCAAGAAATTCACGTTTATTTATGAGCGCCAA 660
 DB 601 CTTTGTGTAGTGAATTTGGGCTTACATCCCAAGAAATTCACGTTTATTTATGAGCGCCAA 660
 QY 661 GTGGAACAAACGAGAGATTTATTCGACTATTCGCTAGAGATGTTATTAATACAGGTCTAAAT 720
 DB 661 GTGGAACAAACGAGAGATTTATTCGATTTATGCGCAAGATGTTATTAATACAGGTCTAAAT 720
 QY 721 AGCTTGAGAGGACAAATGCCCAAGTGTGGTGTGCTTATAATCAATTCGCTAGAGATCTTA 780
 DB 721 AATTTGAGAGGGACAAATGCTGAAAGTTGTTGCGATATAAATCAATTCGCTAGAGATCTTA 780
 QY 781 AGCTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
 DB 781 AGCTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCTGTATTCCA 840
 QY 841 ATAAATACAGTGTCTCAGTTAAACAGGGAAGTTTATACAGAGCGCAATTCGAGCAACAGGG 900
 DB 841 ATGAATACAGTGTCTCAATTAACAGAGAAATTTATACAGATCCCAATTTGGGAGAACAAAT 900
 QY 901 GTAAAT-----ATGGCAAGTATGAATTTGGTATATAATAATGAATGCACCTTCGTTTCCGCT 954
 DB 901 GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAAATAATTAATGCAATTCGTTTTCGCC 960
 QY 955 ATAGAGACTGCGGTATTCCGAGCCCGCATCTACTTTGTTTCTAGAACAACTTACAAT 1014
 DB 961 ATAGAGCTGCGGTATTAGGCTCGCATCTACTTTGTTTCTAGAACAGCTTACAAT 1020
 QY 1015 TTAGCACTTCATCAGATGGAGTGTCTACTAGGCGATATGACTTTACTGGCGGGGACACACA 1074
 DB 1021 TTTAGCGTATTAAATCGATGGAGTAATACTCAATATATGAATTTACTGGGTGGGACATAGA 1080
 QY 1075 ATTCATCTCGGCCAATAGGAGCGGATTAATATCCTCAACGCATGGGTCTACCAATCT 1134
 DB 1081 CTTGAATCGCAACAAATTAAGGGGGTCAATTAAGTACTCTGACACACGGAATACCAATCT 1140
 QY 1135 TCTATTAACTCTGTAAGATTATCTTCTCTCGAGAGCTATATTGGACTGAATCATAT 1194
 DB 1141 TCTATTAACTCTGTAACATTACAGTTTCAATCTCGAGAGCTTTATAGAACAGATCATTT 1200
 QY 1195 GCAGAGTGTCTTATGCGGGAATTTACCTTGAACTTATTCATGGTGTCCCTACTGTAGA 1254
 DB 1201 GCAGGGATAAATAT-----ACTTCTAACTACTCTCTGTGAATGGAGTACTTTGGGCTAGA 1254
 QY 1255 TTTTAAATTTAGGAACCTCGAATACTTTTGAAGAGGTACTGCTAACTATATAGTCAACCC 1314
 DB 1255 TTTTAAATGGAGAAATCCCTCGAATTTCTCTT---AGAGGTAGCCTTCTCTATATAGGG 1311
 QY 1315 TATGAGTCACTGGGCTTCAATTAAGAAATTCAGAAATTCAGAACTGAATTTACCACGAGAAACA 1374
 DB 1312 TATCTGGAGTGGGACACAACCTATTTGATTTCAGAACTGAATTTACCACGAGAAACA 1371
 QY 1375 GAACCAACCAATTTATGAATCATATAGTCTATAGTGTATCTCACAATAGGGCTCATTTCA 1434

Db 1372 GAAAGCCAAATTAATGAATCTTACAGTCATAGATTATCTAAATATAAGACTAATATACAGA 1431
Qy 1435 TCTAGGGTCATGTACCAAGTATATCTTGGAGCGCACCGTAGTGACAGATCGTACAAATACC 1494
Db 1432 AACACTTTGAGAGCACCAAGTATATCTTGGACGCACCGTAGTGACAGATCGTACAAATACC 1491
Qy 1495 ATTAGTTCAGATAGCATNAACAAATACCAATTCGTTGGTAAATCATTCAACTTAAATTCAGGT 1554
Db 1492 ATTAGTTCAGATAGCATNAACAAATACCAATTCGTTGGTAAATCATTCAACTTAAATTCAGGT 1551
Qy 1555 ACCTCTGTAGTCAGTGGCCAGATTTACAGAGAGGGATATAATCCGAACCTAACCGTTAAT 1614
Db 1552 ACCTCTGTAGTCAGTGGCCAGATTTACAGAGAGGGATATAATCCGAACCTAACCGTTAAT 1611
Qy 1615 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATAATACATCATTAACAGCGGTATCCCGTG 1674
Db 1612 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATAATACATCATTAACAGCGGTATCCCGTG 1671
Qy 1675 AGAGTTCGTATGCTGCTTCTCAAAACATGGTCTAGAGGTAACTGTGCGAGGGAGTACT 1734
Db 1672 AGAGTTCGTATGCTGCTTCTCAAAACATGGTCTAGAGGTAACTGTGCGAGGGAGTACT 1731
Qy 1735 ACCTTTGTATCAAGGATCCCTAGTACTATGAGTGCAAATGAGTCTTTGACATCTCAATCA 1794
Db 1732 ACCTTTGTATCAAGGATCCCTAGTACTATGAGTGCAAATGAGTCTTTGACATCTCAATCA 1791
Qy 1795 TTTAGATTTGCAAGATTTCTCTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA 1854
Db 1792 TTTAGATTTGCAAGATTTCTCTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA 1851
Qy 1855 AGTATAAGTAATTAATGAGGTAGACAAACGTTTCTCTTTGATPAAATTTGAATTCATCCA 1914
Db 1852 AGTATAAGTAATTAATGAGGTAGACAAACGTTTCTCTTTGATPAAATTTGAATTCATCCA 1911
Qy 1915 ATTACTGCAACCTCGAAGCAATACGATTTAGAAAGGGCGCAGAGCGGTGATGCT 1974
Db 1912 ATTACTGCAACCTCGAAGCAATACGATTTAGAAAGGGCGCAGAGCGGTGATGCT 1971
Qy 1975 CTGTTTACTAATACGAATCCAAAGAGATTTGAAACAGAGATGTGACAGATTTATCATATTGAT 2034
Db 1972 CTGTTTACTAATACGAATCCAAAGAGATTTGAAACAGAGATTTATCATATTGAT 2031
Qy 2035 CAAATATCCAAATTTAGTGGGTGTTTATCGGATGAATTTCTGTTAGATGAAAGAGAGAA 2094
Db 2032 CAAATATCCAAATTTAGTGGGTGTTTATCGGATGAATTTCTGTTGATGAAAGAGAGAA 2091
Qy 2095 TTACTTGAGAAAGTGAATATGCGAAACGACTCAGTGTGAAGAACTTACTCCAAGT 2154
Db 2092 TTACTTGAGAAAGTGAATATGCGAAACGACTCAGTGTGAAGAACTTACTCCAAGT 2151
Qy 2155 CCAAACTTCAATCCATCAATGAACCAACAGACTTTCATATCTAATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCAATCCATCAATGAACCAACAGACTTTCATATCTAATGAGCAATCGAAT 2211
Qy 2215 TTCAATCTATCCATGAACAAATCTGAACATGGATGGTGGGAGGTGAGAACATTTACAATC 2274
Db 2212 TTCAATCTATCCATGAACAAATCTGAACATGGATGGTGGGAGGTGAGAACATTTACAATC 2271
Qy 2275 CAGGAAGGAATGAGTATTTAAAGAGATTTACGTTCACACTACCGGGACTTTTAAATGAG 2334
Db 2272 CAGGAAGGAATGAGTATTTAAAGAGATTTACGTTCATCTACCGGGACTTTTAAATGAG 2331
Qy 2335 TGTATCCGAGCTATTTATCAAAAAATAGAGAGTCCGAATTTAAAGCTTATACTCGC 2394
Db 2332 TGTATCCGAGCTATTTATCAAAAAATAGAGAGTCCGAATTTAAAGCTTATACTCGC 2391
Qy 2395 TACCAATTAAGAGGTATATTGAAGATAGTCAAGATTTTAGAGATATATTGATTCGTTAT 2454
Db 2392 TACCAATTAAGAGGTATATTGAAGATAGTCAAGATTTTAGAGATATATTGATTCGTTAT 2451
Qy 2455 AATCGGAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGGTTCAGTT 2514
Db 2452 AATCGGAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGGTTCAGTT 2511

Qy 2515 GAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATCGACACCAATTTTGAATGGAT 2574
Db 2512 GAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATCGACACCAATTTTGAATGGAT 2571
Qy 2575 CCTGATCTAGATTTGTTCTCGCAGAGATGGAGAAAATGTGCGCATCATTTCCCATCTTTC 2634
Db 2572 CCTGATCTAGATTTGTTCTCGCAGAGATGGAGAAAATGTGCGCATCATTTCCCATCTTTC 2631
Qy 2635 TCTTTGGATATTTGATTTGAGTGCACACACTTGCATGAGAACTCTAGGCGTGGGTGTA 2694
Db 2632 TCTTTGGATATTTGATTTGAGTGCATGAGAACTCTAGGCGTGGGTGTA 2691
Qy 2695 TTTCAAGATTAAGACGCGAAGAGGTTCATCAAGACTAGGGAATCTGGAATTTTATTCGAAG 2754
Db 2692 TTTCAAGATTAAGACGCGAAGAGGTTCATCAAGACTAGGGAATCTGGAATTTTATTCGAAG 2751
Qy 2755 AAAACCATTTAGAGGAAGCACTGTCTGTGTGAAGAGAGAGAGAGAGAGAGAGAG 2814
Db 2752 AAAACCATTTAGAGGAAGCACTGTCTGTGTGAAGAGAGAGAGAGAGAGAGAGAG 2811
Qy 2815 AAAACGTGAAAACCTACAAATTCGAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2874
Db 2812 AAAACGTGAAAACCTACAAATTCGAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2871
Qy 2875 GATGCTTTATTTAGATTTCTCAATATATAGATTTTCAAGCGGATACAAACATTTGGCATG 2934
Db 2872 GATGCTTTATTTAGATTTCTCAATATATAGATTTTCAAGCGGATACAAACATTTGGCATG 2931
Qy 2935 ATTCAATCGGAGATTAACCTTTGTCATCGAATTCGAGAGGCTTATCTCTCAGAAATTTATCT 2994
Db 2932 ATTCAATCGGAGATTAACCTTTGTCATCGAATTCGAGAGGCTTATCTCTCAGAAATTTATCT 2991
Qy 2995 GTTATCCCGGTGTAAATGCGGAAATTTTTCAGAAATTTAGAGGTCGCAATTTACTCTGCA 3054
Db 2992 GTTATCCCGGTGTAAATGCGGAAATTTTTCAGAAATTTAGAGGTCGCAATTTACTCTGCA 3051
Qy 3055 ATCTCCCTATACGATCGAGAAATGCTGTTAAAAATGCTGATTTTAAATAGGATTTAGCA 3114
Db 3052 ATCTCCCTATACGATCGAGAAATGCTGTTAAAAATGCTGATTTTAAATAGGATTTAGCA 3111
Qy 3115 TGCTGGAATGTAAAGGGCATGTAGTGTACAAAGAGCCATCACCGTTCGTCTCTGTT 3174
Db 3112 TGCTGGAATGTAAAGGGCATGTAGTGTACAAAGAGCCATCACCGTTCGTCTCTGTT 3171
Qy 3175 ATCCAGAAATCGGAGAGCAAGTGTCAACAGCGTTCCGCTCTGTCGGGGCGTGGCTAT 3234
Db 3172 ATCCAGAAATCGGAGAGCAAGTGTCAACAGCGTTCCGCTCTGTCGGGGCGTGGCTAT 3231
Qy 3235 ATCTCTCGGTGTACAGCGGTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAA 3294
Db 3232 ATCTCTCGGTGTACAGCGGTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAA 3291
Qy 3295 ATCGAGAAACATACAGAGCAACTAAAAATTTAAAAATCTGTGAAGAGAGAGAGTGTATCCA 3354
Db 3292 ATCGAGAAACATACAGAGCAACTAAAAATTTAAAAATCTGTGAAGAGAGAGAGTGTATCCA 3351
Qy 3355 ACAGATACAGAACTGTAAATGATTTACTGCACCAAGGTACAGCAGTATGTAATTC 3414
Db 3352 ACAGATACAGAACTGTAAATGATTTACTGCACCAAGGTACAGCAGTATGTAATTC 3411
Qy 3415 CGTAATCTGATATGAGGATGCAATGAACTGTATCTACAGCATCTGTTAAATTTACAAA 3474
Db 3412 CGTAATCTGATATGAGGATGCAATGAACTGTATCTACAGCATCTGTTAAATTTACAAA 3471
Qy 3475 CCGACTTATGAAGAGAAACCGTATACAGATGTACAGAGATTAATTCATTTGGAATATGAC 3534
Db 3472 CCGACTTATGAAGAGAAACCGTATACAGATGTACAGAGATTAATTCATTTGGAATATGAC 3531
Qy 3535 AGAGGATGTGAATTTATCCACCTACCGCTGTTATATGACAAAGAGATTTAGAAATAC 3594
Db 3532 AGAGGATGTGAATTTATCCACCTACCGCTGTTATATGACAAAGAGATTTAGAAATAC 3591

QY	3595	TTCCGAGAACCGATAGGATGATGATTGGAGAACGAGGGAAGTTTATTGTA	3654	Db	61	TGGAATCATTTCCGACAAATGAATCTATCAACCGATGCTCGTATTGAGGATAGCTTGTGT	120
Db	3592	TTCCGAGAACCGATAGGATGATGATTGGAGAACGAGGGAAGTTTATTGTA	3651	QY	121	ATAGCCGAGGGGAATAATATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT	180
QY	3655	GACAGCGTGGAAATTACTCTCTTATGGAGGAATAG	3687	Db	121	ATAGCCGAGGGGAACAATATCGATCCATTTGTGTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	3652	GACAGCGTGGAAATTACTCTCTTATGGAGGAATAG	3684	QY	181	AACATAGCTGGTAGAATACTAGGTGTTATAGCGGTACCGTTTCTGCTGGACAAATAGCTAGT	240
RESULT 4				Db	181	AACATAGCTGGTAGAATACTAGGTGTTATAGCGGTACCGTTTCTGCTGGACAAATAGCTAGT	240
AAA89221				QY	241	TTTTATAGTTTTTCTTGTGTTGAAATTTATGCGCCCGCGGAGAGATCAGTGGGAAATTTTC	300
ID	AAA89221	standard; DNA; 3684 BP.		Db	241	TTTTATAGTTTTTCTTGTGTTGAAATTTATGCGCCCGCGGAGAGATCCTTTGGGAAATTTTC	300
AC	AAA89221;			QY	301	CTAGAACATGTCGAACAACCTTATAATCAACAAATACAGAAATGCTTAGGAATACGGCA	360
DT	19-MAR-2001	(first entry)		Db	301	CTAGAACATGTCGAACAACCTTATAAGACAACAAGTAACAGAAATACTAGGGATACGGCT	360
DE	Bacillus thuringiensis	delta-endotoxin gene 158C2c.		QY	361	CTTGCTCGAATTACAAGTTTACGAGATTCCTTTTAGAGCCTATCAACAGTCACCTTGAAGAT	420
XX	Delta-endotoxin; 158C2c; toxin; lepidoptera; crystal protein;			Db	361	CTTGCTCGAATTACAAGTTTACGAGATTCCTTTTAGAGCCTATCAACAGTCACCTTGAAGAT	420
KW	transgenic plant; insect resistance; crop protection; biological control;			QY	421	TGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTTATACCCAAATATATAGCC	480
KW	ss.			Db	421	TGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTTATACCCAAATATATAGCC	480
OS	Bacillus thuringiensis.			QY	481	TTAGAACTTGATTTTCTTAATCGGATGCCGCTTTTCGCAATTTAGAAACCAAGAGTTCCA	540
XX	US6150589-A.			Db	481	TTAGAACTTGATTTTCTTAATCGGATGCCGCTTTTCGCAATTTAGAAACCAAGAGTTCCA	540
XX	21-NOV-2000.			QY	541	TTATTAAATGTTATATGCTCAAGCTGCMAAATTTTACACCTATTATTATTAGAGAGATGCTCT	600
XX	31-OCT-1997; 97US-00961803.			Db	541	TTATTAAATGTTATATGCTCAAGCTGCMAAATTTTACACCTATTATTATTAGAGAGATGCTCT	600
XX	23-MAY-1995; 95US-00448170.			QY	601	CTTTTTGTAGTGAATTTTGGGCTTTACATCGCAGGAAATTCACAGTTATTATTAGAGCCCAA	660
XX	(MYCO) MYCOGEN CORP.			Db	601	CTTTTTGTAGTGAATTTTGGGCTTTACATCGCAGGAAATTCACAGTTATTATTAGAGCCCAA	660
XX	Narva KE, Stelman S, Payne J, Cummings DA, Cannon RJC;			QY	661	GTGGAAACAACGAGAGATTAATTCGGACTATTTCGGTAGAATGGTATTAATACAGTCTAAAT	720
XX	WPI; 2001-049107/06.			Db	661	GTGGAAACAACGAGAGATTAATTCGTATTAATTCGGCAAGATGGTATTAATACGGGTTTAAAT	720
DR	P-PSDB; AAB19950.			QY	721	AGCTTGAGAGGACAAATGCCGCAAGTTCGGTGGCTTATAAATCAATTCCTGTAGAGATCTA	780
PT	New genes isolated from Bacillus thuringiensis which encode toxins			Db	721	AATTTGAGAGGACAAATGCTGAAGTTGGTTCGGATATAATCAATTCCTGTAGAGACTTA	780
PT	against lepidopteran insects and which can be used to transform various			QY	781	ACGTTAGGGTATTAGATCTAGTGGCACTATTTCCTCAAGCTATGACACTCGCACTTATCCA	840
PT	hosts to express the B.t. toxin and confer insect resistance.			Db	781	ACGTTAGGGTATTAGATCTAGTGGCACTATTTCCTCAAGCTATGACACTCGCACTTATCCA	840
XX	Example 2; Col 19-24; 25pp; English.			QY	841	ATAAATACAGTGTCTCAGTTTACAGGGGAGTTTATACAGACGCAATTCGAGCAACAGG	900
XX	This is the partial nucleotide sequence of the novel delta-endotoxin			Db	841	ATGAATACCAAGTGTCTCAATTAACAAGAGAAATTTATACAGATCCAAATTTGGGAGAACAA	900
CC	1582Cc gene of Bacillus thuringiensis (B.t.) isolate PS158C2 (NRRL B-			QY	901	GTAAAT-----ATGCCAAGTAGTAATTCGGTATATAAATAATGACACTTCGTTTCGGCT	954
CC	19872). To isolate novel toxin genes from PS158C2, total cellular DNA was			Db	901	GCACCTTCAGGATTTGCAAGTAGCAATTCGGTTTATTAATTAATGACACCTCGTTTCTGCC	960
CC	subjected to PCR amplification using primers (see AAA89222-23) based on			QY	955	ATAGAGACTGCGGTTATCCGAAGCCGCACTACTCTTGATTTTCTAGAAACAACCTTAAAT	1014
CC	conserved sequences of B.t. toxins. The PCR products were used as probes			Db	961	ATAGAGGCTGCGGTTATTAGGCTCCGCACTACTTGTATTTTCCAGAACAGACTTACAAAT	1020
CC	to screen a PS158C2 gene library. Plasmid subclone pMYC2383 was obtained			QY	1015	TTTAGCACTTCATCAGATGGAGTGTCTACTAGGCATATCACTACTCGCGGGGACACA	1074
CC	which contained the entire coding sequence for 158C2c (see AAA89221), the			Db	1021	TTTAGCACTTCATCAGATGGAGTGTCTACTAGGCATATCACTACTCGCGGGGACACA	1080
CC	3' portion of the 158C2b gene (see AAA89220) and the 5' portion of 158C2d			QY	1075	ATTCAATCTCGGCCAATAGGAGCGGATTAATACTCAACGATCGGTCTTACCAATACT	1134
CC	(see AAA89218). The 158C2a gene was isolated on a separate subclone,			Db	1081	CTTGAATCCGGAACAATAAGGGGGTCATTAGTACTCTCGACACACGGAATACCAATACT	1140
CC	pMYC2387. These novel genes encode delta-endotoxins (see AAB19947-50)			QY	1135	TCTATTAATCTCTGTAGAGATTATCAATCTCTCTCGAGAGCTATATTGGACTGAATCATAT	1194
CC	that are active against lepidopteran pests. Expression in transformed			Db	1141	TCTATTAATCTCTGTAGAGATTATCAATCTCTCTCGAGAGCTTATATAGAACAGAAATCATTT	1200
CC	plants confers resistance to such pests, while transformed microorganisms						
CC	expressing the delta-endotoxins can be applied to sites where they will						
CC	proliferate and be ingested by the pest						
SQ	Sequence 3684 BP; 1232 A; 628 C; 803 G; 1021 T; 0 U; 0 Other;						
Query Match	88.6%; Score 3266.6; DB 4; Length 3684;						
Best Local Similarity	93.4%; Pred. No. 0;						
Matches 3449; Conservative	0; Mismatches 229; Indels 15; Gaps 3;						
QY	1	TTGACTTCAATAGGAAAATGAGATGAATTAATAAATGCTTTATCGATTCCAGCTGTA	60				
Db	1	TTGACTTCAATAGGAAAATGAGATGAATTAATAAATGCTTTATCGATTCCAGCTGTA	60				
QY	61	TCGAATCATTTCCACAAATGGATCTATCAACGATGCTCGTATTGAGGATTCCTTGTGT	120				

QY	1195	GCAGGAGTGCCTTCTATGCGGGAATTTACCTTGAACCTATTCTCATGTGTCCCTACTGTGTAGA	1254
DB	1201	GCAGGGAATAATAT-----ACTTCTAACTACTCTGTGAATGAGTACCTTGGCGCTAGA	1254
QY	1255	TTTAATTTTAGGAACCTTCAGAACTACTTTTGAAGAGAGTACTGCTAACTATAGTCAACCC	1314
DB	1255	TTTAAATTGGAGAAATCCCTGAAATTCCTT---AGAGTGACCTTCTCTATACTATAGGG	1311
QY	1315	TATGAGTCACTCGGCGTTCAAATTAAGATTCAGAAACTGAAATTCACACAGAGAAACAACA	1374
DB	1312	TATACTGGAGTGGGACACAACTATTTGATTCAGAAACTGAAATTCACACAGAGAAACAACA	1371
QY	1375	GAAAGCAACCAATTAATGAATCATATAGTCATAGGTTATCTCACATAGGGCTCATTTACAA	1434
DB	1372	GAAAGCAACCAATTAATGAATCTTACAGTCATAGATTTATCTAATTAAGACATAATATCAGGA	1431
QY	1435	TCTAGGTCATGTACACAGTATATCTTGGACGCACCGTGTAGTCAGATCGTACAAATATACC	1494
DB	1432	AACACTTTGAGAGCACCAAGTATATCTTGGACGCACCGTGTAGTCAGATCGTACAAATATACC	1491
QY	1495	ATTAGTTCAGATAGCATAAACAATAACCAATTCGTGTAAATCATTTCAACCTTAATTCAGGT	1554
DB	1492	ATTAGTTCAGATAGCATAAACAATAACCAATTCGTGTAAATCATTTCAACCTTAATTCAGGT	1551
QY	1555	ACCTCTGTAGTCAGTGGCCAGGAATTTACAGGAGGGATATATCCGAACTAAGCTTTAAT	1614
DB	1552	ACCTCTGTAGTCAGTGGCCAGGAATTTACAGGAGGGATATATCCGAACTAAGCTTTAAT	1611
QY	1615	GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATAATACATCATTCACGCGGTATCGCGTG	1674
DB	1612	GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATAATACATCATTCACGCGGTATCGCGTG	1671
QY	1675	AGAGTTCGTTATGCTGCTTCTCAAAACAATGGTCTCGAGGGTAACTGTCGAGGGAGTACT	1734
DB	1672	AGAGTTCGTTATGCTGCTTCTCAAAACAATGGTCTCGAGGGTAACTGTCGAGGGAGTACT	1731
QY	1735	ACTTTTCATCAAGGATTCCTCTAGTACTATGAGTGCAAATGAGTCTTTTGAATCTCAATCA	1794
DB	1732	ACTTTTCATCAAGGATTCCTCTAGTACTATGAGTGCAAATGAGTCTTTTGAATCTCAATCA	1791
QY	1795	TTTAGATTTGCAAGATTTCTCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATA	1854
DB	1792	TTTAGATTTGCAAGATTTCTCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATA	1851
QY	1855	AGTATAGTAATAATGCAAGTAGACAAACCGTTTCACCTTTGATATAAATGAAATTCATTCCA	1914
DB	1852	AGTATAGTAATAATGCAAGTAGACAAACCGTTTCACCTTTGATATAAATGAAATTCATTCCA	1911
QY	1915	ATTACTGCACCTTCGAAGCAGAGATACGATTTAGAAAGGCGCAAGAGGGCGGTGAATGCT	1974
DB	1912	ATTACTGCACCTTCGAAGCAGAGATACGATTTAGAAAGGCGCAAGAGGGCGGTGAATGCT	1971
QY	1975	CTGTTTACTAATACGAATCCAAGAGAGATTGAAAAACAGATGTGACAGATTTATCATATTGAT	2034
DB	1972	CTGTTTACTAATACGAATCCAAGAGAGATTGAAAAACAGATGTGACAGATTTATCATATTGAT	2031
QY	2035	CAAGTATCCAAATTTAGTGGCGTGTATTTCGGATGAATCTGCTTAGATGAAAGAGAGAGAA	2094
DB	2032	GAAGTATCCAAATTTAGTGGCGTGTATTTCGGATGAATCTGCTTTGGATGAAAGAGAGAA	2091
QY	2095	TTACTTCGAGAAAGTAAATATGCCAAACGACTCAGTGCATGAAAGAACTTTACTCCAGAT	2154
DB	2092	TTACTTCGAGAAAGTAAATATGCCAAACGACTCAGTGCATGAAAGAACTTTACTCCAGAT	2151
QY	2155	CCAAACTTCACATCCATCAATTAAGCAACCCAGACTTCTATCTACTAATGAGCAATCGAAT	2214
DB	2152	CCAAACTTCACATCCATCAATTAAGCAACCCAGACTTCTATCTACTAATGAGCAATCGAAT	2211
QY	2215	TTGACATCTATCCATGAAACAACTGGAACATGGATGGTGGGGAAGTGTAGAACATTACATC	2274
DB	2212	TTGACATCTATCCATGAAACAACTGGAACATGGATGGTGGGGAAGTGTAGAACATTACATC	2271

QY	2275	CAGGAAGAAATGACGTATTTAAAGAGAAATTA	CGTCACTACCGGGACCTTTTAATGAG	2334
DB	2272	CAGGAAGAAATGACGTATTTAAAGAGAAATTA	CGTCACTACCGGGACCTTTTAATGAG	2331
QY	2335	TGTTATCCGAGTATTTATATCAAAANATAGGAGAGT	CGGAAATTAAGAGCTTATACCTCGC	2394
DB	2332	TGTTATCCGAGTATTTATATCAAAANATAGGAGAGT	CGGAAATTAAGAGCTTATACCTCGC	2391
QY	2395	TACCAATTAAGAGGGTATATTGAAGATAGTCAAGAT	TTAGAGATATATTTGAATTCGTTAT	2454
DB	2392	TACCAATTAAGTGGCTATATTGAAGATAGTCAAGAT	TTAGAGATATATTTGAATTCGTTAT	2451
QY	2455	AATCGGAAAATGAAAATTCGATGTTCCAGGTACCGAGT	CCGTAATGCGCGCTTTCAGTT	2514
DB	2452	AATCGGAAAATGAAAATTCGATGTTCCAGGTACCGAGT	CCGTAATGCGCGCTTTCAGTT	2511
QY	2515	GAAGGCCAATCGGAAGTGGAGAACCGAATCGATCGCAC	CCACATTTTGAATTCGAAAT	2574
DB	2512	GAAGGCCAATCGGAAGTGGAGAACCGAATCGATCGCAC	CCACATTTTGAATTCGAAAT	2571
QY	2575	CCTGATCTAGATGTTCTTCGAGAGATGGAGAAAAAT	TGCGGCATCATTTCCCATCATTTTC	2634
DB	2572	CCTGATCTAGATGTTCTTCGAGAGATGGAGAAAAAT	TGCGGCATCATTTCCCATCATTTTC	2631
QY	2635	TCCTTTGGATATTTGATATTTGGATGCGACAGACTT	GGCATGAGAAATTCAGCGGTGGTGTT	2694
DB	2632	TCCTTTGGATATTTGATATTTGGATGCGACAGACTT	GGCATGAGAAATTCAGCGGTGGTGTT	2691
QY	2695	TTCAAGATTTAAGACGAGGAGGTCTATGCAAGACTAGG	GAATCTGGNAATTTATTTGAAGAG	2754
DB	2692	TTCAAGATTTAAGACGAGGAGGTCTATGCAAGACTAGG	GAATCTGGNAATTTATTTGAAGAG	2751
QY	2755	AAACCATATTTAGGAGAAAGCACTGCTCGTGTGGAAG	AGACAGAGAAAAATTCGAGAGAC	2814
DB	2752	AAACCATATTTAGGAGAAAGCACTGCTCGTGTGGAAG	AGACAGAGAAAAATTCGAGAGAC	2811
QY	2815	AAACGTGAAAAATCTAATTTGGAAAAAACAAGAGTAT	ATACAGAGGCAAAAGAGCTGTG	2874
DB	2812	AAACGTGAAAAATCTAATTTGGAAAAAACAAGAGTAT	ATACAGAGGCAAAAGAGCTGTG	2871
QY	2875	GATCGTTTATTTGATAGTCTCAATATATAGATTAACA	GGCATCAACATTCGCGCATG	2934
DB	2872	GATCGTTTATTTGATAGTCTCAATATATAGATTAACA	GGCATCAACATTCGCGCATG	2931
QY	2935	ATTCATCGCGCAGATAAATCTGTTTCATCGAAATTCG	AGAGCTTATCTGTGAGAAATTTATCT	2994
DB	2932	ATTCATCGCGCAGATAAATCTGTTTCATCGAAATTCG	AGAGCTTATCTGTGAGAAATTTATCT	2991
QY	2995	GTATCCCGGTGTPAAATGCGGAAATTTTGAAGAAAT	TAGAAGGTCCCATTTACCTGCA	3054
DB	2992	GTATCCCGGTGTPAAATGCGGAAATTTTGAAGAAAT	TAGAAGGTCCCATTTACCTGCA	3051
QY	3055	ATCTCCCTATACGATGCGAGAAATGTCGTTTAAAAAT	TGGTGATTTTAATTAAGATGACCA	3114
DB	3052	ATCTCCCTATACGATGCGAGAAATGTCGTTTAAAAAT	TGGTGATTTTAATTAAGATGACCA	3111
QY	3115	TGCTGGAAATGTAAAAAGGCGATGTAGATGTACACAG	AGCCATCACCGTTCCTGCTTGT	3174
DB	3112	TGCTGGAAATGTAAAAAGGCGATGTAGATGTACACAG	AGCCATCACCGTTCCTGCTTGT	3171
QY	3175	ATCCAGAAATGGGAAGCAGAAAGTGTCAACAGCAGTT	CGCGTCTGTCCGGGGCGTGCTAT	3234
DB	3172	ATCCAGAAATGGGAAGCAGAAAGTGTCAACAGCAGTT	CGCGTCTGTCCGGGGCGTGCTAT	3231
QY	3235	ATCTCCGTGTCAACGGTACAAAGGGGATATGGAGAGG	TTGTGTGTAAAGATCCATGAA	3294
DB	3232	ATCTCCGTGTCAACGGTACAAAGGGGATATGGAGAGG	TTGTGTGTAAAGATCCATGAA	3291
QY	3295	ATCGAGAACCAATACAGACGAACCTAAAAATTTAAAA	AACTGTGGAAGAGGAGGTATCCA	3354
DB	3292	ATCGAGAACCAATACAGACGAACCTAAAAATTTAAAA	AACTGTGGAAGAGGAGGTATCCA	3351
QY	3355	ACGGATACAGGAACGCTGTAAATGATTTACTGTCACAC	CAAGGTTACACAGCATGTATTC	3414

Db 3352 ACGGATACAGGACGTGTAATGATTATCTGCACCAAGGTACAGCAGCATGTAATTC 3411
Qy 3415 CGTAATGCTGGATGATGAGGATGCAATGAAAGTTGATCTACAGCATCTGTTAATTACAA 3474
Db 3412 CGTAATGCTGGATGATGAGGATGCAATGAAAGTTGATCTACAGCATCTGTTAATTACAA 3471
Qy 3475 CCGACTTATGAGAGAAACGTATACAGATGTCGAAAGAGATAATCATTTGTAATATGAC 3534
Db 3472 CCGACTTATGAGAGAAACGTATACAGATGTCGAAAGAGATAATCATTTGTAATATGAC 3531
Qy 3535 AGAGGGTATGTAATTTACCACTACCACTGCTGTTATATGACAAAAGTAATGAGATAC 3594
Db 3532 AGAGGGTATGTAATTTACCACTACCACTGCTGTTATATGACAAAAGTAATGAGATAC 3591
Qy 3595 TTCCAGAAACCGGATAGGATGATGAGATTGGAGAAACCGAAGGGAAGTTTATTGTA 3654
Db 3592 TTCCAGAAACCGGATAGGATGATGAGATTGGAGAAACCGAAGGGAAGTTTATTGTA 3651
Qy 3655 GACAGCGTGAATTTACTCTTATGGAGGATAG 3687
Db 3652 GACAGCGTGAATTTACTCTTATGGAGGATAG 3684

RESULT 5
AAS02489
ID AAS02489 standard; DNA; 3684 BP.

AC AAS02489;

XX 29-AUG-2001 (first entry)

XX B. thuringiensis DNA encoding a toxic crystal protein, CryET54.

XX Delta endotoxin; Lepidopteran-active; crystal protein; insecticide;
XX transgenic plant; corn; wheat; soybean; oat; cotton; rice; rye; sorghum;
XX sugarcane; tomato; tobacco; kapok; flax; potato; barley; turf grass;
XX pasture grass; berry; fruit; legume; vegetable; ornamental plant; shrub;
XX cactus; tree cell; gypsy moth; looper; tobacco budworm; spruce budworm;
XX cotton leaf perforator; CryET54; ds.

XX Bacillus thuringiensis.

XX Key Location/Qualifiers
FH 1. .3684
FT CDS /*tag= a
FT /*product= "CryET54"
FT /partial
FT /note= "No start codon"

XX WO200119859-A2.

XX 22-MAR-2001.

XX 13-SEP-2000; 2000WO-US025361.

XX 15-SEP-1999; 99US-0153995P.

XX (MONS) MONSANTO CO.

XX Baum JA, Chu C, Donovan WP, Gilmer AJ, Rupar MJ;

XX WPI; 2001-281518/29.

XX P-PSDB; AAU02046.

XX Lepidopteran-active Bacillus thuringiensis delta-endotoxin polypeptides
XX and the polynucleotides that encode them, useful for increasing the
XX insect resistance of plant.

XX Claim 17; Page 168-169; 173pp; English.

XX The sequence encodes a B. thuringiensis Lepidopteran-active delta-
XX endotoxin, crystal protein CryET54. The Lepidopteran-active B.

CC thuringiensis delta-endotoxin polypeptides may be used as compositions
CC that are applied to plant crops to protect them from insect damage. The
CC polynucleotides may be used in the production of transgenic plants that
CC express the insecticidal polypeptides and consequently have improved
CC insect resistance compared to non-transformed plants. Monocotyledonous or
CC dicotyledonous plants may be protected in this way, for example corn,
CC wheat, soybean, oat, cotton, rice, rye, sorghum, sugarcane, tomato,
CC tobacco, kapok, flax, potato, barley, turf grass, pasture grass, berry,
CC fruit, legume, vegetable, ornamental plant, shrub, cactus and/or tree
CC cell. A wide range of insects (e.g. gypsy moth, looper, tobacco budworm,
CC cotton leaf perforator and spruce budworm) may be affected by application
CC of the insecticidal polypeptides (full details given in specification).
CC Note: The present sequence does not have a cryET number assigned
CC explicitly in the specification, cryET54 (see table 5, page 56) is not
CC assigned to any Seq.ID number, the index has matched the spare sequence
CC to the spare cryET number, which may be incorrect

XX Seq Sequence 3684 BP; 1232 A; 626 C; 803 G; 1023 T; 0 U; 0 Other;

Query Match 88.4%; Score 3260.2; DB 4; Length 3684;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

Qy 1 TTGACTTCAATAGGAAATAATGAGATGAAATTAATAATGCTTTATCGATTCCAGCTGTA 60

Db 1 TTGACTTCAATAGGAAATAATGAGATGAAATTAATAATGCTTTATCGATTCCAGCTGTA 60

Qy 61 TCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTGAGGATTCTTTGTGT 120

Db 61 TCGAATCATTTCCGACAAATGAATCTATCAACCGATGCTCGTATTGAGGATAGTTGTGT 120

Qy 121 ATAGCGGGGGGAATAATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATT 180

Db 121 ATAGCGGGGGGAACAATATCGATCCATTTGTAGCGCATCAACAGTCCAAACGGGTATT 180

Qy 181 AACATAGCTGGTAGAATACTAGGTGCTATTAGCGGTACCGTTTCTGGGACAAATAGTACT 240

Db 181 AACATAGCTGGTAGAATACTAGGTGCTATTAGCGGTACCGTTTCTGGGACAAATAGTACT 240

Qy 241 TTTTATAGTTTTCTTTGGTGAATATTGGCCCCCGGCGAGAGATCAGTGGGAAATTTTC 300

Db 241 TTTTATAGTTTTCTTTGGTGAATATTGGCCCCCGGCGAGAGATCCTTTGGGAAATTTTC 300

Qy 301 CTAGAACATGTCGAACACTTATAATCAACAAATAACAGAAANTGCTAGGATACGGCA 360

Db 301 CTAGAACATGTCGAACACTTATAAGAACAAACAGTAACAGAAATACTAGGATACGGCT 360

Qy 361 CTTGCTCGATTACAAAGTTTAGGAGATTCTTTAGAGCTTATCAACAGTCACTTGAAGAT 420

Db 361 CTTGCTCGATTACAAAGTTTAGGAAATTTCTTTAGAGCTTATCAACAGTCACTTGAAGAT 420

Qy 421 TGGCTAGAAAACCGGTGATGCAAGAACGAGAGAGTGTCTTTATACCAATATATAGCC 480

Db 421 TGGCTAGAAAACCGGTGATGCAAGAACGAGAGAGTGTCTTTATACCAATATATAGCC 480

Qy 481 TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTTCGCAATTAGAAACCAAGAGTTCCA 540

Db 481 TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTTCGCAATTAGAAACCAAGAGTTCCA 540

Qy 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTACACCTATTATTATTGAGAGATGCCTCT 600

Db 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTATACCTATTATTATTGAGAGATGCCTCT 600

Qy 601 CTTTTTGTAGTGAATTTGGGCTTACATCGCAGGAAATTTCAACGTTATTATGAGCGCCAA 660

Db 601 CTTTTTGTAGTGAATTTGGGCTTACATCCCAAGAAATTTCAACGTTATTATGAGCGCCAA 660

Qy 661 GTGGAACAAACGAGAGATTATTCGAGCTATTTCGCTAGAAATGGTATAATACAGGCTAAAT 720

Db 661 GTGGAACAAACGAGAGATAATTCGATTATTTCGCAAGATGGTATAATACGSGTTAAAT 720

Qy 721 AGCTTGAGAGGACAAATGCCGAAGTTGGGTGGTTATTAATCAATTCGGTAGAGATCTA 780

Db 721 AGCTTGAGAGGACAAATGCCGAAGTTGGGTGGTTATTAATCAATTCGGTAGAGATCTA 780

Db 721 AATTGAGAGGGAATAAAGTCTGAAAGTGTGGTTCGGATATAAATCAATTCCTAGAGACTTA 780
Qy 781 AGCTTAGGGGTATATAGATCTAGTGGCACTATTTCCCAAGCTATGACACACTCGCACTTATCCA 840
Db 781 ACCTAGAGATATATAGATCTAGTGGCACTATTTCCCAAGCTATGACACACTCGGTATCCA 840
Qy 841 ATAAATACAGATGCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
Db 841 ATGAATACCAAGTCTCAATTAACAGAGAAATTTATACAGATCCCAATTTGGAGAAACAAAT 900
Qy 901 GTAAAT-----ATGCCAGTATGATTTGGTATATAATAATGCAACCTTCGTTTTCCGCT 954
Db 901 GCACCTTCAGAGATTTCGAAGTACGAATTTGGTTTAAATAATATGCAACCTTCGTTTTCTGCC 960
Qy 955 ATAGAGACTCGCGTTATCCGAAGCCCGCATCTACTTGTATTTCTAGAACCAACTTACAAT 1014
Db 961 ATAGAGGCTCGGTTATTTAGGCTCGCATCTACTTGTATTTCCAGAACAGCTTACAAT 1020
Qy 1015 TTTAGACATTCATCAGATGAGTGTCTACTAGGCATATGACTTACTTGGCGGGGCACACA 1074
Db 1021 TTTAGCGTATTAAGTCGATGAGTAATACTCAATATATGAATTACTGGGTGGGACATAGA 1080
Qy 1075 ATTCAATCTCGGCCAATAGAGGCGGATTAATAATACCTCAACGATGGGTCTACCAATACT 1134
Db 1081 CTTGAATCCGSAACAATAAGGGGGTCAATTAAGTACCTGGGACACACGGAATACCAATACT 1140
Qy 1135 TCTATTAATCTGTGAAGTATCATTTCTCTCGAGACGTATATTTGGACTGAATCATAT 1194
Db 1141 TCTATTAATCTGTGAAGTATCATTTCTCGAGACGTATATTTGGACTGAATCATAT 1200
Qy 1195 GCAGGATGCTTCTATGCGGAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTGAGA 1254
Db 1201 GCAGGATAAATAT-----ACTTCTAACTACTCTCTGATGGAGTACCTTTGGGCTAGA 1254
Qy 1255 TTTAATTTTAGAACCTCAGAAATCTTTTGAAGAGGTACTGTCTAACTATAGTCAACCC 1314
Db 1255 TTTAATTTGAGAAATCCCGTGAATCTCTTT---AGAGGTAGCTTTCTCTATACTATAGG 1311
Qy 1315 TATGAGTCACTCGGGCTTCAATTTAAAGATTCAGAAATTCAGAAATACCAACAGAAACA 1374
Db 1312 TATCTGGAGTGGGACACACTATTTGATTCAGAACTGAATTTACCACAGAAACA 1371
Qy 1375 GAAACGACAAATATGAATCATATAGTCAATAGTTTATCTCATAGGGCTCATTTTCAAA 1434
Db 1372 GAAACGACAAATATGAATCTTACAGTCAATAGTTTATCTAATATGAAGTCAATATACGA 1431
Qy 1435 TCTAGGTGCTATGACAGTATTTCTTGGACGACCGGTAGTGCAGATCGTACAAATACC 1494
Db 1432 AACACTTTGAGAGCACCAATATTTCTTGGACGACCGGTAGTGCAGATCGTACAAATACC 1491
Qy 1495 ATTAGTTCAGATAGCATACAAATACCATTTGGTAAATCATTCACCTTAATTCAGGT 1554
Db 1492 ATTAGTTCAGATAGCATACAAATACCATTTGGTAAATCATTCACCTTAATTCAGGT 1551
Qy 1555 ACCTCTGTAGTCACTGGCCAGGATTTACAGAGGGGATATATTCGGAATCAACCTTAAT 1614
Db 1552 ACCTCTGTAGTCACTGGCCAGGATTTACAGAGGGGATATATTCGGAATCAACCTTAAT 1611
Qy 1615 GGTAGTGTACTAAGTATGGTCTTAATTTTAAATATACATNTTACAGGGTATCGCGTG 1674
Db 1612 GGTAGTGTACTAAGTATGGTCTTAATTTTAAATATACATNTTACAGGGTATCGCGTG 1671
Qy 1675 AGAGTTTGTATGCTGCTTCTCAAAATGCTGAGGGTAACTGTCGGAGGAGTACT 1734
Db 1672 AGAGTTTGTATGCTGCTTCTCAAAATGCTGAGGGTAACTGTCGGAGGAGTACT 1731
Qy 1735 ACTTTTGTACAGGATTCCTAGTACTATGAGTGCAAATGAGTCTTTTGACATCTCAATCA 1794
Db 1732 ACTTTTGTACAGGATTCCTAGTACTATGAGTGCAAATGAGTCTTTTGACATCTCAATCA 1791
Qy 1795 TTTAGATTTGAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGTGGAATA 1854
Db 1792 TTTAGATTTGAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGTGGAATA 1851

Qy 1855 AGTATAAGTAAATAATGACAGGTAGACAAACGTTTTCATTTTGATAAAATTTGAATTCATTCCA 1914
Db 1852 AGTATAAGTAAATAATGACAGGTAGACAAACGTTTTCATTTTGATAAAATTTGAATTCATTCCA 1911
Qy 1915 ATTACTCGAACTTCCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGCGCGTGAATGCT 1974
Db 1912 ATTACTCGAACTTCCGAAGCAGAAATATGATTTAGAAAGGGCGCAAGCGCGTGAATGCT 1971
Qy 1975 CTGTTTACTAATACGAATCCGAAGAGATTGAACAGAGATGTGACAGATTTATCATTTGAT 2034
Db 1972 CTGTTTACTAATACGAATCCGAAGAGATTGAACAGAGATGTGACAGATTTATCATTTGAT 2031
Qy 2035 CAAATATCCAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTAGATGAAAAGAGAGAA 2094
Db 2032 GAAATATCCAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTGATGAAAAGAGAGAA 2091
Qy 2095 TTAATTTGAGAAAGTGAATATTCGAAACGACTCTAGTATGAAAGAACTTTACTCCAAGAT 2154
Db 2092 TTAATTTGAGAAAGTGAATATTCGAAACGACTCTAGTATGAAAGAACTTTACTCCAAGAT 2151
Qy 2155 CCAAACTTCAATCCATCAATAAGCAACGAGCTTCAATCTACTAATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCAATCCATCAATAAGCAACGAGCTTCAATCTACTAATGAGCAATCGAAT 2211
Qy 2215 TTTCACTCTATCCATGAACAAATCTGAACATGATGGTGGGAGTGAAGCAATTAACAATC 2274
Db 2212 TTTCACTCTATCCATGAACAAATCTGAACATGATGGTGGGAGTGAAGCAATTAACAATC 2271
Qy 2275 CAGGAAGAAATGACGTTATTTAAAGAGAAATTAACGTACACTACCGGGGACTTTTAAATGAG 2334
Db 2272 CAGGAAGAAATGACGTTATTTAAAGAGAAATTAACGTACACTACCGGGGACTTTTAAATGAG 2331
Qy 2335 TGTATTCGAGCTATTTATATCAAAAAATAGAGAGTCCGAATTTAAAGCTTATATCTCGC 2394
Db 2332 TGTATTCGAGCTATTTATATCAAAAAATAGAGAGTCCGAATTTAAAGCTTATATCTCGC 2391
Qy 2395 TACCAATTAAGAGGGTATATTTAAAGAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2454
Db 2392 TACCAATTAAGAGGGTATATTTAAAGAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2451
Qy 2455 AATCGAAACATGAAACATTTGATTCGAGATACCGGATCCGATGATGCGCGCTTTCACTT 2514
Db 2452 AATCGAAACATGAAACATTTGATTCGAGATACCGGATCCGATGATGCGCGCTTTCACTT 2511
Qy 2515 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATGCGCACCACTTTTGAATGGAAT 2574
Db 2512 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATGCGCACCACTTTTGAATGGAAT 2571
Qy 2575 CTTGATCTAGATTTGTTCTCTGAGAGATGGAAGAAATGTCGCGATCATTTCCCATCATTTTC 2634
Db 2572 CTTGATCTAGATTTGTTCTCTGAGAGATGGAAGAAATGTCGCGATCATTTCCCATCATTTTC 2631
Qy 2635 TCTTTGGATATTTGATTTGGAATGCAACGACTTGCATGAGAAATCTAGGCGTGTGGTGGTA 2694
Db 2632 TCTTTGGATATTTGATTTGGAATGCAACGACTTGCATGAGAAATCTAGGCGTGTGGTGGTA 2691
Qy 2695 TTTCAAGATTAAGACGCGAGAAAGGTCATGCAAGACTAGGGAATCTGGAATTTTATTTGAAGAG 2754
Db 2692 TTTCAAGATTAAGACGCGAGAAAGGTCATGCAAGACTAGGGAATCTGGAATTTTATTTGAAGAG 2751
Qy 2755 AAAACATTTATTTAGGAGAGCACTGTCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAC 2814
Db 2752 AAAACATTTATTTAGGAGAGCACTGTCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAC 2811
Qy 2815 AAAAGTGAAGAACTCAATTTGGAAGAAACAAACGAGTATATACAGAGGCAAGAGAGAGAGAGAG 2874
Db 2812 AAAAGTGAAGAACTCAATTTGGAAGAAACAAACGAGTATATACAGAGGCAAGAGAGAGAGAGAG 2871
Qy 2875 GATGCTTTATTTGATGATTTCTCAATATATATGATTTACAGCGGATACAAACATTTGGCGATG 2934
Db 2872 GATGCTTTATTTGATGATTTCTCAATATATGATTTACAGCGGATACAAACATTTGGCGATG 2931

Db 711 TTTATGATGGTATGCTCAAGCTGCAAAATTTACACCTATTATATGAGATGCTCT 770
Qy 601 CTTTTTGGTAGTAAATTTGGGCTTACATCGCAGGAAATTCACGTTATATGAGCGCCAA 660
Db 771 CTTTTTGGTAGTAAATTTGGGCTTACATCGCAGGAAATTCACGTTATATGAGCGCCAA 830
Qy 661 GTGGAACAAACGAGAGATTTATCCGACTATTCGCTAGATGCTATAATACAGGTCTAAAT 720
Db 831 GTGGAACAAACGAGAGATTTATCCGACTATTCGCTAGATGCTATAATACAGGTCTAAAT 890
Qy 721 AGCTTGAGAGGACAAATCCCGCAAGTTGGGTCGCTTATATCAATTCGCTAGAGATCTA 780
Db 891 AGCTTGAGAGGACAAATCCCGCAAGTTGGGTCGCTTATATCAATTCGCTAGAGATCTA 950
Qy 781 AGCTTGAGGCTATTAGATCTAGTGCACCTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
Db 951 AGCTTGAGGCTATTAGATCTAGTGCACCTATTTCCCAAGCTATGACACTCGCACTTATCCA 1010
Qy 841 ATAAATACGAGTCTCAGTTAAACAGGGAAGTTTATACAGCGCAATTTGGAGCAACAGGG 900
Db 1011 ATAAATACGAGTCTCAGTTAAACAGGGAAGTTTATACAGCGCAATTTGGAGCAACAGGG 1070
Qy 901 GTAAATATGCGAAGTATGAATTTGGTATATAATTAATGCACTTCTGCTATAGAG 960
Db 1071 GTAAATATGCGAAGTATGAATTTGGTATATAATTAATGCACTTCTGCTATAGAG 1130
Qy 961 ACTGCGGTTATCCGAAGCCGCTACTTCTGAGATTTCTAGAACAACTTACAAATTTTACG 1020
Db 1131 GCTGCGGCTATCCGAAGCCGCTACTTCTGAGATTTCTAGAACAACTTACAAATTTTACG 1190
Qy 1021 ACTTCATACGATGGAGTCTACTAGGCATATGACTTACTGGCGGGGCAACAAATTCAA 1080
Db 1191 GCTTCATACGATGGAGTAAATACAGGCATATGACTTATTTGGCGGGGCAACAGATTCAA 1250
Qy 1081 TCTCGGCAATAGGAGGGGATTAATACCTCAACGATGGGTCTACCAATCTTCTATTT 1140
Db 1251 TCTCGGCAATAGGAGGGGATTAATACCTCAACGATGGGTCTACCAATCTTCTATTT 1310
Qy 1141 AATCTGTAAGATTTACTTCTCTCGAGACGTATATGCACTGCAATCATATGACGGA 1200
Db 1311 AATCTGTAAGATTTACTGCTCGACATCTCGAGACGTATATGCACTGCAATCATATGACGGA 1370
Qy 1201 GTGCTTCTATGGGGAATTTACCTTGAACTTCTCATGGTGTCCCTACTGTTAGATTTAAT 1260
Db 1371 GTGCTTCTATGGGGAATTTACCTTGAACTTCTCATGGTGTCCCTACTGTTAGGTTAAT 1430
Qy 1261 TTTAGGAACCTTCAGAAATCTTTTGAAGAGGTACTGCTAATAGTCAACCTTATGAG 1320
Db 1431 TTTAGGAACCTTCAGAAATCTTTCTGATAGAGGTACCGCTAATAGTCAACCTTATGAG 1490
Qy 1321 TCACCTGGGCTTCAATTTAAAGATTCAGAACTGAAATACCACAGAAACCAACAGACGA 1380
Db 1491 TCACCTGGGCTTCAATTTAAAGATTCAGAACTGAAATACCACAGAAACCAACAGACGA 1550
Qy 1381 CCAATTTATGATCATATAGTCAATAGGTTATCTCATAGGCTCATTTTCAAACTTAGG 1440
Db 1551 CCAATTTATGATCATATAGTCAATAGGTTATCTCATAGGTTATTTTACAACTCCAGG 1610
Qy 1441 GTGCAATGACAGATATTTCTTGGAGCGACCGTATGAGATCGTACAAATACCAATAGT 1500
Db 1611 GTGCAATGACAGATATTTCTTGGAGCGACCGTATGAGATCGTACAAATACCAATAGT 1670
Qy 1501 TCAGATAGCATACCAATACCAATTTGTTAAATCATTTCAACCTTAAATTCAGGTACCTCT 1560
Db 1671 CCAATTAGAATCACCACCAATCCAAATGGTAAAGCATCCGAACTTCTCAAGGTACCACT 1730
Qy 1561 GTAGTCAGTGGCCAGGATTTACAGGAGGGGATATAATCCGAACTTAAATGAGT 1620
Db 1731 GTTGTAGAGGACAGGATTTACTGCTGGGATATTTCTTGGAGAAACCAATACTGGTGA 1790
Qy 1621 GTACTAAGTGGGTCTTAAATTTTAAATATACATCATTTACAGCGGTATCGCGTGAAGTT 1680

Db 1791 TTTGACCGATAAGAGTAACTGTTAAACGGAACATTTAAACAAAGATATCGTAGGATTC 1850
Qy 1681 CGTTATGCTGCTTCTTCAACCAATGCTCTGAGGGTAACTGTCGAGGGAGTACTACTTTT 1740
Db 1851 CGCTATGCTTCAACTGTAGATTTTGAATTTCTTTGTATCACGTGGAGGTACTCTGTAAT 1910
Qy 1741 GATCAAGGATTTCCCTTAGTACTATGAGTGCAAAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
Db 1911 AATTTAGATTTCTTACGTACATGACAGTGGAGCAAGCAATATAATACGGAATTTTGTG 1970
Qy 1801 TTTGAGAAATTTCTTGTAGGTATTTAGTGCATCTGCGAGTCAA---ACTGCTGGAATAAGT 1857
Db 1971 AGACGTGCTTTTACTACACCTTTTACTTTTACAAATTTCAAGATATAATTTGCAACGCT 2030
Qy 1858 ATAAAGTATAATGACAGGTAGACAAACGTTTCACTTTGATAAAATGGAATTCATTTCCAAT 1917
Db 2031 ATTCAAGCCCTTAGTGGAATTTGGGAGTGTATATAGATAAAATTTGAATTTATTCAGTT 2090
Qy 1918 ACTGCAACCTTTCGAAGCAGAAATACGATTTAGAAAGGCGCAAGGCGGTGAAATGCTCTG 1977
Db 2091 ACTGCAACCTTTCGAAGCAGAAATATGATTTAGAAAGGCGCAAGGCGGTGAAATGCTCTG 2150
Qy 1978 TTTACTAATACGATCCGAAGAAAGATTGAAACAGATGTGACAGATTCATATTCATCAA 2037
Db 2151 TTTACTAATACGATCCGAAGAAAGATTGAAACAGATGTGACAGATTCATATTCATCAA 2210
Qy 2038 GTATCCAAATTTAGTGGGCTTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2097
Db 2211 GTATCCAAATTTAGTGGGCTTTTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAAATTA 2270
Qy 2098 CTTGAGAAAGTGAATATATCGAAACGACTCAGTGATGAAAGAAACCTTACTCCAGATCCA 2157
Db 2271 CTTGAGAAAGTGAATATATCGAAACGACTCAGTGATGAAAGAAACCTTACTCCAGATCCA 2330
Qy 2158 AACTTCATCCATCAATTAAGCAACCGACTTCTATCTACTAATAGAGCAATTCGAATTC 2217
Db 2331 AACTTCATCCATCAATTAAGCAACCGACTTCTATCTACTAATAGAGCAATTCGAATTC 2390
Qy 2218 ACATCTATCCATGAACCAATCTGAAACATGATGGGGAAGTGAGAACTTACATCCAG 2277
Db 2391 ACATCTATCCATGAACCAATCTGAAACATGATGGGGAAGTGAGAACTTACATCCAG 2450
Qy 2278 GAAGGAATACGATTTTAAAGAGAAATTTACGTCACACTACCGGGGACTTTTAAATGAGTGT 2337
Db 2451 GAAGGAATACGATTTTAAAGAGAAATTTACGTCACACTACCGGGGACTTTTAAATGAGTGT 2510
Qy 2338 TATCCGAGTATTTATATCAAAAATAGGAGAGTCGGAATTTAAAGCTTTATCTCGCTAC 2397
Db 2511 TATCCGAGTATTTATATCAAAAATAGGAGAGTCGGAATTTAAAGCTTTATCTCGCTAC 2570
Qy 2398 CAATTAAGAGGATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAT 2457
Db 2571 CAATTAAGAGGATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAT 2630
Qy 2458 GCGAAACATGAAACATTTGGATGTTCCAGGTAACCGAGTCCGATGCGGCTTTTCAAGTGA 2517
Db 2631 GCGAAACATGAAACATTTGGATGTTCCAGGTAACCGAGTCCGATGCGGCTTTTCAAGTGA 2690
Qy 2518 AGCCCAATCGGAAGGTGCGGAGAACCGAATCGATGCGGACCACTTTTGAATGGAATCCT 2577
Db 2691 AGCCCAATCGGAAGGTGCGGAGAACCGAATCGATGCGGACCACTTTTGAATGGAATCCT 2750
Qy 2578 GATCTAGATTTCTTCTCGAGAGATGGGAAATTTGCGCATCTATCCCATCATTTCTCT 2637
Db 2751 GATCTAGATTTCTTCTCGAGAGATGGGAAATTTGCGCATCTATTTCCCATCATTTCTCT 2810
Qy 2638 TTTGATATTTGATATTTGATGACAGACTTTCATGAGAACTTAGCGGTGCGGTGATTC 2697
Db 2811 TTTGATATTTGATATTTGATGACAGACTTTCATGAGAACTTAGCGGTGCGGTGATTC 2870
Qy 2698 AAGATTAAGACGAGGAGGTCTGCAAGACTAGGAACTTGGAAATTTTATTTGAAGAGAAA 2757
Db 2871 AAGATTAAGACGAGGAGGTCTGCAAGACTAGGAACTTGGAAATTTTATTTGAAGAGAAA 2930

Db 427 ATTGCTCGATTAGAAGTCTAGAGAGAGGCTATAGATCTTACAGCAGGCTCTTGAAACT 486
Qy 421 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCAAATATATAGCC 480
Db 487 TGGTTAGATAACCGGAAATGATGCAAGATCAAGAGCAATTTATCTTGAGCGCTATGTGCT 546
Qy 481 TTAGAACTTGATTTCTTAATCGGATGCGCTTTTCGCAATTTAGAAAACCAAGAAAGTTCCA 540
Db 547 TTAGAACTTTGACATTACTACTGCTATACCGCTTTTCAGAAATCGAAATGAAGAGTTCCA 606
Qy 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTAATTAATGAGAGATCCCTCT 600
Db 607 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTAATTAATGAGAGCGCATCC 666
Qy 601 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACAGTTTATTAATGAGCGCCAA 660
Db 667 CTTTTTGGTAGTGAATTTGGGAGTGGCATCTCCGATGTTAACCAATATTACCAGAACAA 726
Qy 661 GTGGAAACAAACGAGAGATTTTCCGACTATTCGGTAGAATGGTATAATACAGGTCTAAAT 720
Db 727 ATCAGATATACAGAGGAATTTCTAACCAATTCGCTACAAATGGTATAATACAGGGCTAAAT 786
Qy 721 ASCITTAGAGGGACAAATCCCAAGTTGGGTGCGTTAATCAATTCGCTAGAGATCTA 780
Db 787 AACTTAAGAGGGACAAATCTGAAAGTTGGTTGCGGTATAATCAATTCGCTAGAGACCTA 846
Qy 781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
Db 847 ACGTTAGGGGTATTAGATTTAGTGGCTATTTCCCAAGCTATGACTCGCACTTATCCA 906
Qy 841 ATAATAACAGTGTCTCAGTTAAACAGGGAAAGTTTATACAGCGCAATTTGGAGCAACAGGG 900
Db 907 ATCAATACAGAGTGTCTCAGTTAAACAGAGAAAATTTATACAGATCCCAATTTGGGAGAACAAAT 966
Qy 901 GTAAAT-----ATGCGAGTATGAATGGTATATAATTAATGCACTTCGTTTTCCGCT 954
Db 967 GCACCTTCAGGATTTGCAAGTACGAATGGTTTAAATAATGCACTTCGTTTTCCGCT 1026
Qy 955 ATAGAGACTCGGTTTATCGAAGCCGCACTTACTTACTTGTATTTCTAGAACAACTTACAAT 1014
Db 1027 ATAGAGGCTGCCATTTTCAGGCTCCGCACTTACTTGTATTTCCAGAACAACTTACAAT 1086
Qy 1015 TTTAGCACTTCAATCAAGTAGAGTGTACTAGGCAATAGACTTACTGCGGGGGCACACA 1074
Db 1087 TACAGTGCATCAAGCGCTTGGAGTAGCACTCAACATATGAATATTGCGTGGGACATAGG 1146
Qy 1075 ATTCATCTCGGCCAATAGGAGCGGATTAATATCTCAACGCAATGGGTCTACCA---AT 1131
Db 1147 CTTAACCTCCGCCCAATAGGAGGACATTAATATCTCAACAAAGGACTTACTAATAAT 1206
Qy 1132 ACTTCTATTAAATCCTGTAAGATTATCATTTCTCTCGAGACGTATATTGGACTGAATCA 1191
Db 1207 ACTTCANATTATCTGTAAATACATCAGTTTAGCTCTCGAGAGCTTTTATAGAACAGATCA 1266
Qy 1192 TATCAGGAGTGTCTCTATGGGGAATTTACCTTGAACCTTATTTCATGGTGTCCCTACTGTT 1251
Db 1267 AATGAGGGACAAATAT-----ACTATTACTCTCTGTGAATGGAGTACCTTGGGCT 1320
Qy 1252 AGATTTAATTTTAGAAACCTCAGAAATACTTTTGAAGAGGTAATGCTTAACTATAGTCAA 1311
Db 1321 AGATTTAATTTTATAAACCTCAGAAATATTATGAAGAGGGCGGCATACCTACAGTCAA 1380
Qy 1312 CCCTATGAGTCACTCGGGCTTCAATTTAAAGATTCAGAACTGAAATTAACCAACAGAAACA 1371
Db 1381 CGGTATCAGGAGTTGGGATTCATTTATTTGATTCAGAACTGAATTTACCACAGAAACA 1440
Qy 1372 ACAGAACGACCAAAATTTGAATCATATAGTCAATAGGTTATCTCATAGGGCTCATTTTCA 1431
Db 1441 ACAGAACGACCAAAATTTGAATCATATAGTCAATAGTATCTCATATAGGACTAATCATA 1500
Qy 1432 CAATCTAGGTCATGTACAGTATATTTCTGGACGACCGGTAGTGCAGATGCTGACAAAT 1491
Db 1501 GGAAACACTTTGAGAGCACCACTTATTTCTTGGACGCACTCGTAGTGCAGATCGTACGAAT 1560

Qy 1492 ACCATTAGTTCAGATAGCATAAACAAATACCAATTTGGTAAAAATCAATTCAAACCTTAATCCA 1551
Db 1561 ACGATTGACCAAAATAGAAATTTACAAATACCAATTTGGTAAAAAGCACTGAATCTTCATCCA 1620
Qy 1552 GGTACTCTGTAGTCAAGTGGCCAGGATTTACAGAGGGGATATAATCCGAACTAACGTT 1611
Db 1621 GGTGTACTGTGTGGAGGGCCAGGATTTACAGGTGGGGATATCTTTGTAAGAACAAAT 1680
Qy 1612 AATGTAAGTGTACTAAGTATGGTCTTAATTTTAAATAATACATCATTAACAGCGGTATCGC 1671
Db 1681 ACGGTTACATTTGGAGATATACGATTAATTAATGTGCCATTAATCCAAAGATATCGC 1740
Qy 1672 GTGAGAGTTCTGTATGCTCTTCTCAAAACAAATGGTCTCGAGGGTAACTGTCCGAGGGAGT 1731
Db 1741 GTAAGGATTCGTATGCTTCTTACTACAGATTTACAATTTTTCACGAGAAATTAATCGAAC 1800
Qy 1732 ACTACTTTGATCAAGGATTTCCCTAGTACTATGAGTGCAAAATGAGTCTTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATTGGTAAATTTCTCAAGAACTATGAATAGGGGGATAATTTTGAATATAGA 1860
Qy 1792 TCATTTAGATTTGAGAAATTTCTGTAGTATTTAGTGCATCTGGCAGTCAAACTCTGGA 1851
Db 1861 AGTTTGAACCTGAGGATTTAGTACTCTTTTAAATTTTAAATGCCCCAAGCAATTC 1920
Qy 1852 ATAAAGTATAAGTAAATATGACAGGTAGACAAAACGTTTCACTTTGTATAAAAATTTGAATTCAT 1911
Db 1921 ACATTTGGGTGCTCAGAGTTTTCAAAATCAGGAAGTTTATATAGATAGAGTTCGAATTTGTT 1980
Qy 1912 CCAATTAATCTGCAACCTTCGAAAGCAATACGATTTAGAAAAGGGCCAAAGAGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAAACATTTGAGGCGAGATATGATTTAGAAAGAGCACAAAAGCGGTGAAT 2040
Qy 1972 GCTCTGTTTACTAATACGAATCCAGAGATTTGAAACAGATGTGACAGATATCATATTT 2031
Db 2041 GCTCTGTTTACTTCTTACAAATCCAGAGATTTGAAACAGATGTGACAGATATCATATTT 2100
Qy 2032 GATCAAGTATCCAAATTTTGTGCGGTGTTTATCGGATGAATTTCTGCTTAGATGAAAAGAGA 2091
Db 2101 GACCAAGTGTCCATATGTTGCGCATGTTTATCAGATGAATTTTGTCTGATGAGAGGGA 2160
Qy 2092 GAATTAATCTGAGAAAGTGAATATGCGAAAACGACTCAGTGTAGTGAAGAAAATTTACTCAA 2151
Db 2161 GAATTAATCTGAGAAAGTGAATATGCGAAGCGACTCAGTGTAGTGAAGAAAATTTACTCAA 2220
Qy 2152 GATCCAACTTCAATCCATCAATAAGCAACAGACTTCATATCTACTAATGAGCAATCG 2211
Db 2221 GATCCAACTTCAATCCATCAATCAGTGGGCAATTTAAGTTTTCGATCCATCGTGGACAATCA 2280
Qy 2212 AATTTCACTCTATCCATGAACAAATCTGAACATGGATGGTGGGAGGTGAGAACATTTACA 2271
Db 2281 AACTTCCCTCTATTAATGAGCTATCTGAACTGATGGTGGGAGGTGCGAATGTTAC 2340
Qy 2272 ATCCAGGAAGAAATGACGTAATTTAAAGAGAAATTCGTCACACTACCGGGGACTTTTAAAT 2331
Db 2341 ATTCCAGGAAGGAAATGACGTAATTTAAAGAGAAATTTACGTCACACTACCGGTACTTTTAAAT 2400
Qy 2332 GAGTGTATCCGAGTATTTATCAAAAATAGAGAGTCCGGAATTTAAAGCTTATATCT 2391
Db 2401 GAGTGTATCCAAATTTATATATCAAAAAATAGAGAGTCCAGAAATTTAAAGCTTATATACG 2460
Qy 2392 CGCTTACCAATTAAGAGGGTATATTGAAGATAGTCAAGATTTTAGAGATATATTTGATTCGT 2451
Db 2461 CGCTTACCAATTAAGAGGGTATATTGAAGATAGTCAAGATCTAGAGATTTATTTAATTCGT 2520
Qy 2452 TATAATCGAAACATGAACAAATTCGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCA 2511
Db 2521 TACAATCGAAAGCATGAACAAATTCGATGTTCCAGGTACCGATTCCTATGCGCGCTTTCA 2580
Qy 2512 GTTGAAGCCCAATCGGAGGTGGGAGAACCGGAATCGATGCGCACCAATTTTGAATGG 2571
Db 2581 GTTGAAGCCCAATCGGAGGTGGGAGAACCAAAATCGATGCGCACCAATTTTGAATGG 2640

```
QY 2572 AATCCTGATCTAGATTGTTCTCCAGAGATGGAGAAAATGTGCGCATCATTTCCCATCAT 2631
Db |||||||
QY 2641 AATCCTGATCTAGATTGTTCTCCAGAGATGGAGAAAATGTGCGCATCATTTCCCATCAT 2700
Db |||||||
QY 2632 TTCTCTTTGGATATTGATATTGGATGACAGACTTTCATGAGAAATCTTAGCGGTGTGGGTG 2691
Db |||||||
QY 2701 TTCATTGGATATTGATTTGGTTCACAGACTTTCATGAGAACTTAGCGGTGTGGGTG 2760
Db |||||||
QY 2692 GTATTCAAGATTAAAGCGCAGGAAGGTTCATGCAAGACTTAGGAACTCTGGAATTTATTGAA 2751
Db |||||||
QY 2761 GTATTCAAGATTAAAGCGCAGGAAGGTTCATGCAAGACTTAGGAACTCTGGAATTTATTGAA 2820
Db |||||||
QY 2752 GAGAAACCATTTATGAGAGAGACACTGCTCTCGTGTGAAGACGACAGAGAAAATCGAGA 2811
Db |||||||
QY 2821 GAGAAACCATTTATGGAAGACACTGCTCTCGTGTGAAGAGCGGAGAAAATCGAGA 2880
Db |||||||
QY 2812 GACAAACGCGTGAATACTTACAAATTTGGAACCAAAACGAGTATATACAGAGCAAAAGAGCT 2871
Db |||||||
QY 2881 GACAAACGCGGAAAACCTTACAAATTTGGAACCAAAACGAGTATATACAGAGCAAAAGAGCT 2940
Db |||||||
QY 2872 GTGATGCTTTATTGATAGATTCTCAATATATAATAGATTACAAGCGGATACAAACATTTGGC 2931
Db |||||||
QY 2941 GTGATGCTTTATTGATAGATTCTCAATATATAATAGATTACAAGCGGATACAAACATTTGGC 3000
Db |||||||
QY 2932 ATGATTTCATCGCGCAGATAAATTTGTTTCATCGAATTCGAGAGCTTATCTGTCAGAAATTA 2991
Db |||||||
QY 3001 ATGATTTCATCGCGCAGATAAATTTGTTTCATCGAATTCGAGAGCGGTATCTTTGAGAAATTA 3060
Db |||||||
QY 2992 TCTGTTATCCCGGTGTAAATGCGGAAAATTTTGAAGAATTAGAAGTTCGCAATTTACCT 3051
Db |||||||
QY 3061 CCTGTTATCCCGGTGTAAATGCGGAAAATTTTGAAGAATTAGAAGTTCACATTTACCT 3120
Db |||||||
QY 3052 GCAATCTCCCTATACGATCGGAGAAATGTCGTTAAATAATGTTGATTTTAATATGATTA 3111
Db |||||||
QY 3121 GCAATGTCCTTATACGATCGGAGAAATGTCGTTAAATAATGTTGATTTTAATATGATTA 3180
Db |||||||
QY 3112 GCATGCTGGAATGTAAAGGCGCATGTAGATGTACAAGAGCCATCACCGTTCGTCCTT 3171
Db |||||||
QY 3181 ACATGTTTGGAAATGTAAAGGCGCATGTAGATGTACAAGAGCCATCATCTTCTGACCTT 3240
Db |||||||
QY 3172 GTTATCCAGAAATGGGAAGCAGAAAGTGTCAAAAGCAGATTTCGCTCTGTCCGGGCGGTGGC 3231
Db |||||||
QY 3241 GTTATCCAGAAATGGGAAGCAGAAAGTGTCAAAAGCAGATTTCGCTCTGTCCGGGCGGTGGC 3300
Db |||||||
QY 3232 TATATCTCTCGTGTACAGCGGTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCAT 3291
Db |||||||
QY 3301 TATATCTCTCGTGTACAGCGGTACAAAGAGGATATGGAGAGGCTGCGTAAACGATCCAT 3360
Db |||||||
QY 3292 GAAATCGAGAAACATACAGACGAACTTAAATTTTAAATACTGTGAAGAGGAGGAGTGTAT 3351
Db |||||||
QY 3361 GAAATCGAGAAACATACAGACGAACTTAAATTTTAAATACTGTGAAGAGGAGGAGTGTAT 3420
Db |||||||
QY 3352 CCAACCGGATACAGAAACGTTGATTAATTAATCTGCAACCAAGGTACAGCAGTATGTAAT 3411
Db |||||||
QY 3421 CCAACCGGATACAGAAACGTTGATTAATTAATCTGCAACCAAGGTACAGCAGTATGTAAT 3480
Db |||||||
QY 3412 TCCGTAATCTCGATATGAGGATGCAATTAAGATTGATACAGCATCTGTTAATTAC 3471
Db |||||||
QY 3481 TCCGTAATCTCGATATGAGGATGCAATTAAGATTGATACAGCATCTGTTAATTAC 3540
Db |||||||
QY 3472 AAACCGACTTATGAAGAGAAACCTATACAGATGTACGAAGAGATTAATCAATCTGTAATAT 3531
Db |||||||
QY 3541 AAACCGACTTATGAAGAGAAACCTATACAGATGTACGAAGAGATTAATCAATCTGTAATAT 3600
Db |||||||
QY 3532 GACAGAGGGTATGTGAATTTATCCACCACTACAGCTGGTTTATATGACAAAGAAATTAGAA 3591
Db |||||||
QY 3601 GACAGAGGGTATGTGAATTTATCCACCACTACAGCTGGTTTATGTGACAAAGAAATTAGAA 3660
Db |||||||
QY 3592 TACTTCCAGAAACCGTATAGGATGTGAGTTGGAGAAACGAGGAGGATTTATT 3651
Db |||||||
QY 3661 TACTTCCAGAAACAGATACAGTATGAGTTGGAGATTGGAGAAACGAGGAGGATTTATT 3720
Db |||||||
QY 3652 GTAGACAGCGTGGAAATTAATCTCTTATGAGGAATAG 3687
```

```
Db 3721 GTAGATAGCTGGAACTACTCTCTCATGGAAATAG 3756
|||||
RESULT 8
AAZ09160
ID AAZ09160 standard; DNA; 3932 BP.
XX AAZ09160;
AC AAZ09160;
DT 18-OCT-1999 (first entry)
XX
DE B. thuringiensis cryET5 toxin DNA.
XX
KW Toxin; cryET4; cryET5; insecticidal; Lepidoptera; transformed plant;
  crystal protein; insect; ds.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FT CDS 67..3756
FT FT /*tag= a
FT FT /product= "cryET5"
FT FT /note= "No ATG start codon"
XX
PN US5942658-A.
XX
PD 24-AUG-1999.
XX
PF 24-JUN-1997; 97US-00881340.
XX
PR 29-JUL-1993; 93US-00100709.
PR 30-DEC-1993; 93US-00176865.
PR 07-JUN-1995; 95US-00474038.
XX
PA (MONS ) MONSANTO CO.
XX
PI Gonzalez JM, Jany CS, Tan Y, Donovan WP;
XX
DR WPI; 1999-493544/41.
DR P-PSDB; AAY30923.
XX
PT Transformed plant comprising insecticidal crystal proteins.
XX
PS Claim 4; Fig 2A-J; 50pp; English.
XX
CC This invention describes novel transformed plants containing Bacillus or
  thuringiensis strain EG5847 insecticidal crystal protein genes cryET4 or
  cryET5. cryET4 and cryET5 are novel toxin genes which produce
  insecticidal proteins with activity against a broad spectrum of insects
  of the order lepidoptera. This sequence encodes the cryET5 protein
  described in the method of the invention
XX
SQ Sequence 3932 BP; 1347 A; 650 C; 840 G; 1095 T; 0 U; 0 Other;
  Query Match 72.9%; Score 2687.2; DB 2; Length 3932;
  Best Local Similarity 83.5%; Pred.No. 0;
  Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;
  1 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATTTATAAATGCTTTATCGATTCCAGCTGTA 60
  67 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATTTATAAATGCTTTATCGATTCCACGGTA 136
  61 TCGAATCATTCACACAAATGGATCTATCACCGAGTCTCGTATTGAGGATTTCTTTGTGT 120
  127 TCGAATCTCTCCACGCAATGAATCTATCACGAGATGCTCGTATTGAAGATAGCTTTGTGT 186
  121 ATAGCCGAGGGAATAATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
  187 GTAGCCGAGGTGAACAAATATGATCCATTTGTTAGCGCATCAACAGTCCAAACGGGTATA 246
  181 AACATAGCTGGTAGAATACTAGGTGTTATTAGGCGTACCGTTTGTCTGGACAAATAGCTAGT 240
  |||||||
```

Db 247 AACATAGCTGTAGAAATATTGGGCGCTATTAGGTGTGCGGTTTGTGCGACAACCTAGCTAGT 306
Qy 241 TTTTATAGTTTCTTGTGGTGAATATTATGGCCCGCGCAGAGATCACTGGGGAATTTTC 300
Db 307 TTTTATAGTTTCTTGTGGGGAATATTATGGCCCTAGTGGCAGAGATCCATGGGAATTTTC 366
Qy 301 CTAGAACATGTGCGAACCACTTATAAATCAACAATAACAGAAATAGCTAGGAATACGGCA 360
Db 367 CTGGAACATGTAGAACCACTTATAAGACAAACAAGTAAACAGAAAAATCTAGGAATACGGCT 426
Qy 361 CTTCGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCTTATCAACAGCTCACTTTGAAGAT 420
Db 427 ATTGCTCGATTAGAAGGCTTAGGAAGAGGCTATAGATCTTACCAGAGGCTCTTTGAAACT 486
Qy 421 TGGCTAGAAAACCGGTGATGATGCAAGAACGAGAGGTTTCTTTATACCCAATATATAGCC 480
Db 487 TGGTTAGATACCGGAATGATGCAAGATCAAGAGCATTTCTTGAGGGCTATGTTGCT 546
Qy 481 TTAGAACTTGATTTTCTTAATGCGGATGCGCTTTTTCGCAATTAGAAAACAGAAAGTTCCA 540
Db 547 TTAGAACTTGATTTTCTTAATGCGGATGCGCTTTTTCGCAATTAGAAAACAGAAAGTTCCA 606
Qy 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTCAGAGATGCCCTCT 600
Db 607 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTCAGAGATGCCATCC 666
Qy 601 CTTTTTGTAGTGAATTTGGGCTTACATCGCAGAGAAATTCACAGTTTATATGAGCGCCA 660
Db 667 CTTTTTGTAGTGAATTTGGGCTTACATCGCAGAGAAATTCACAGTTTATATGAGCGCCA 726
Qy 661 GTGGAACAAACGAGAGATTTATTCGACTATTGCGTGAATGGTATAATACAGGCTTAAAT 720
Db 727 ATCAGATATACAGAGGAATTTCTAACCATTTGCGTACAAATGGTATAATACAGGCGCTAAT 786
Qy 721 AGCTTGAGAGGACAAATCCCGCAGATTTGGTGGCTTATATCAATTCCTGAGATCTA 780
Db 787 AACTTAAGAGGGACAAATGCTGAAAGTTGGTTGGCGTATATCAATTCCTGAGAGACCTA 846
Qy 781 ACGTTAGGGGTATTAGATCTAGTGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 840
Db 847 ACGTTAGGGGTATTAGATCTAGTGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 906
Qy 841 ATAAATACGAGTCTCAGTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
Db 907 ATCAATACGAGTCTCAGTTAAACAGGGAATTTTATACAGATCCAAATTTGGGAGAACAAAT 966
Qy 901 GTAAAT-----ATGGCAAGTATGAATTTGGTATATAATATATGACACTTTGCTCGCT 954
Db 967 GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTATATAATGCAACCACTCGTTTCTGCC 1026
Qy 955 ATAGAGACTGGGTTTATCCGAAGCCCGCATCTACTTGAATTTCTAGAACAACTTACAAAT 1014
Db 1027 ATAGAGGCTGCCATTTTTCAGGCTCCGCATCTACTTGAATTTTCCAGAACAACTTACAAAT 1086
Qy 1015 TTTAGCACTTCATCAACGATGGAGTCTACTAGGCATATGACTTACTGCGGGGGGCACACA 1074
Db 1087 TACAGTGATCAAGCCGTTGGAGTAGCACTCAACATATGAATTTATTTGGTGGGACATAGG 1146
Qy 1075 ATTCAATCTCGGCCAATAGAGGCGGATTAATATCACTCAACGATGGGTCTACCA---AT 1131
Db 1147 CTTAACCTTCCGCCCAATAGAGGGGACATTAATAATACCTCAACACAGGACTTTACTAATAAT 1206
Qy 1132 ACTTCTAATTAATCCGTGAAGTATATCAATCTTCTCGAGAGCTATATTGACACTGAAATCA 1191
Db 1207 ACTTCAATTAATCCGTGAAGTATATCAATCTTCTCGAGAGCTATATTGACACTGAAATCA 1266
Qy 1192 TATGAGGAGTGTCTTATAGGGGAATTTTACCTTGAACTTATTCATGGTGTCCCTACTGTT 1251
Db 1267 AATGAGGAGCAAAATAT-----ACTATTACTACTCTCTGTGATGGAGTACTTGGGCT 1320
Qy 1252 AGATTTAATTTAGGAACCCCTCAGAAATCTTTTGAAGAGGTACTGCTTAACATATAGTCAA 1311
Db 1321 AGATTTAATTTTATAAACCCCTCAGAAATTTTATGAAAGAGGCGGCACCTACCTACAGTCAA 1380

Qy 1312 CCCTATAGTCACTCGGCTTCAATTTAAAGATTTAGAAATCTGAATTAACCAACAGAAACA 1371
Db 1381 CCGTATCAGGAGTTGGGATTCATTTATTTGATTCAGAACTGAATTAACCAACAGAAACA 1440
Qy 1372 ACAGAACGACCAAAATTTAGAAATCATATAGTCAATAGGTTATCTCATAGGGCTCATTTCA 1431
Db 1441 ACAGAACGACCAAAATTTAGAAATCATATAGTCAATAGTCAATAGTATCTCATATAGGACTAATCATA 1500
Qy 1432 CAATCTAGGGTGCATGTACCAAGTATATCTTTGGAGCGACCCGTAGTGCAGATCGTCAAAAT 1491
Db 1501 GGAACACTTTTGAGAGCACCAAGTCTATTCTTTGGAGCGCATCTGTAGTGCAGATCGTACGAAT 1560
Qy 1492 ACCATTAGTTCAGATAGCATAAACAATAACATCTGGTAAATCAATTCAACTTTAAATTTCA 1551
Db 1561 ACGATTGGAACCAATAGAAATTTACACAATACCAATTTGGTAAAGCACTGAATCTTTCATTTCA 1620
Qy 1552 GGTACCTCTGTAGTCACTGGGCCAGGATTTTACAGAGGGGATATAATCCGAACTAAACGTT 1611
Db 1621 GGTGTACTGTGTTGGAGGGCCAGGATTTTACAGGTGGGATATCTCTCGTAGAACAAAT 1680
Qy 1612 AATGCTAGTGTACTAAGTATAGGTCCTTAATTTTAAATATACATCATTTACAGGGTATCGC 1671
Db 1681 ACGGTTACATTTGGAGATATACGAATTAATTAATTTAGTCCATTTATCCCAAGATATCGC 1740
Qy 1672 GTGAGAGTTCGTTATGCTGCTTCTCAACAATGCTCCTGAGGGTAACTGTCCGAGGAGT 1731
Db 1741 GTAAGGATTCGTTATGCTTCTACTACAGATTTACAATTTTTCACAGAAATTAATGGAACC 1800
Qy 1732 ACTACTTTTGTATCAAGGATTCCTTAGTACTATAGTGCATAATAGTCTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATGTTAAATTTCTCAAGAACTATGATATAGGGGGGATTAATTTAGAAATATAGA 1860
Qy 1792 TCATTTAGTTTGCAGAAATTTCTGTAGTATTTAGTGCATCTGSCAGTCAACCTGCTGGA 1851
Db 1861 AGTTTGAAGCTGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCCAACACATTC 1920
Qy 1852 ATAAGTATAAGTATAATGAGGTAGCAACAGTTTCTCACTTTGATTAATTTGAATTTCAAT 1911
Db 1921 ACATTTGGTGTCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGTGCATTTGTT 1980
Qy 1912 CCAATTAAGTCAACCTTTGAGAGCAAGATACGATTTAGAAAGGGCGCAAGCGCGTGAAT 1971
Db 1981 CCAGCAGAGGTAAACATTTGAGGCGAAGATATGATTTAGAAAGAGCACAAAGCGCGTGAAT 2040
Qy 1972 GCTCTGTTTACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTTATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTGAAACAGATGTGACAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTTCGGATGAATTTCTGTTAGATGAAAGAGA 2091
Db 2101 GACCAAGTGTCCAATATGTTGGGCAATGTTTATCAGATGAATTTTGTCTGGATGAGAGCGA 2160
Qy 2092 GAATTTACTTGAGAAAGTGAATATTCGAAACGACTCAGTGTATGAAAGAAACTTACTCCAA 2151
Db 2161 GAATTTTGGAGAAAGTGAATATTCGAAACGACTCAGTGTATGAAAGAAACTTACTCCAA 2220
Qy 2152 GATCCAACTTCACTCATCAATTAAGCAACCACTTATCATATCTATTAATGAGCAATCG 2211
Db 2221 GATCCAACTTCACTCATCACTCAGTGGGCAATTAAGTTTTCGATCCATCGATGGAACATCA 2280
Qy 2212 AATTTTCATCTATTCATGAACCAATCTGAACATGATGGTGGGAGGTGAGAACATTTACA 2271
Db 2281 AACTTCCCTCTATTAAATGAGCTATCTGAAACATGATGGTGGGAGGTGCGAATGTTTACC 2340
Qy 2272 ATCCAGGAAGGAATGAGTATTTTAAAGAGAAATTTACGTCACACTACCGGGGACTTTTAAAT 2331
Db 2341 ATTCCAGGAAGGATGAGTATTTTAAAGAGAAATTTACGTCACACTACCGGGGACTTTTAAAT 2400
Qy 2332 GAGTGTATTCGAGCTATTTATATCAAAAAATAGGAGAGTGGGAATTTTAAAGCTTTTACT 2391
Db 2401 GAGTGTATTCGAATTTTATATCAAAAAATAGGAGAGTGGGAATTTTAAAGCTTTTACTACG 2460

Db 67 TTGACTTCAAAATAGGAAATAAGAGAAATGAAATATAAATGCTTTTATCGATTCCAAAGGTA 126
Qy 61 TCGAATCATTTCCACAAATGATCTATCACCAGATGCTCGTATTGAGGATCTTTTGTT 120
Db 127 TCGAATCCTTCCACGCAATGATCTATCACCAGATGCTCGTATTGAGGATGCTTGTT 186
Qy 121 ATAGCGGAGGGAATAATCAATCACTGTTAGGCAATCAACAGTCCAAACGGGTATT 180
Db 187 GTAGCGGAGGTGAACAATATTGATCCATTTGTTAGGCAATCAACAGTCCAAACGGGTATA 246
Qy 181 AACATAGCTGTAGAAATCACTAGTGTATTAGGCGTACCGTTCCTGGACCAATAGCTAGT 240
Db 247 AACATAGCTGTAGAAATCACTAGTGTATTAGGCGTTCCTGGACCAATAGCTAGT 306
Qy 241 TTTTATAGTTTCTTGTGGTGAATATAGGCCCCCGGCGAGAGATCAGTGGGAAATTTTC 300
Db 307 TTTTATAGTTTCTTGTGGGAAATATAGGCCCCCGGCGAGAGATCAGTGGGAAATTTTC 366
Qy 301 CTAGAACATGTGCAACAACTTATAATCAACAAATAACAGAAATGCTAGGAAATACGGCA 360
Db 367 CTGGAACATGTAGAACAACTTATAAGACAAACAAAGTAAACAGAAATACGGCT 426
Qy 361 CTGCTCGATTACAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
Db 427 ATTGCTCGATTAGAAAGGCTAGGAAAGAGGCTATAGATCTTACAGCAGGCTCTTGAAACT 486
Qy 421 TGGCTAGAAACCGTGATGATCAAGAACGAGAAGTGTCTTTATACCCCAATATATAGCC 480
Db 487 TGGTTAGATTAACGAAATGATGCAAGATCAAGAGCAATTTCTTGAGGCTATGTTGCT 546
Qy 481 TTAGAACTTGATTTCTTAATCGATGCGCTTTTCGCAATTTAGAAACCAAGAGTTCCA 540
Db 547 TTAGAACTTGATTTCTTAATCGATGCGCTTTTCGCAATTTAGAAACCAAGAGTTCCA 606
Qy 541 TTATTAATGGTATGCTCAAGCTGCAAAATTTACACCTATTATTTAGAGATGCTCT 600
Db 607 TTATTAATGGTATGCTCAAGCTGCAAAATTTACACCTATTATTTAGAGAGACGATCC 666
Qy 601 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTTCAAGTTATTATGAGGCGCAA 660
Db 667 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTTCAAGTTATTATGAGGCGCAA 726
Qy 661 GTGGAACAAACGAGAGATTTATTCGACTATTGCGTAGAATGATTAATACAGGTCATAAT 720
Db 727 ATCAGATATACAGAGAAATTTCTAACCATTTGCGTACAAATGATTAATACAGGCGTAAAT 786
Qy 721 AGCTTGAGAGGACAAATGCGCAAGTTGGGCGTATTAATCAATTCGCTAGAGATCTA 780
Db 787 AACTTAAGAGGACAAATGCTGAAAGTTGGTTGCGGTATAATCAATTCGCTAGAGACCTA 846
Qy 781 ACGTTAGGGGTATTAGATCTAGTGCACCTATTCCCAAGCTATGACACTCGCACTTTATCCA 840
Db 847 ACGTTAGGGGTATTAGATTTAGTAGCCCTATTCCCAAGCTATGATCTCGCACTTTATCCA 906
Qy 841 ATAAATACGAGTGTCTCAGTTAAACAGGGAAGTTTATACAGAGCGCAATTTGGAGCAACAGGG 900
Db 907 ATCAATACGAGTGTCTCAGTTAAACAGGGAAGTTTATACAGATCCAAATTTGGGAGACAAAT 966
Qy 901 GTAAAT-----ATGGCAAGTATGAATTTGGGTATTAATTAATGACACCTTCGTTTCCGCT 954
Db 967 GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAAATAATTAATGACACCACTCGTTTCTGCC 1026
Qy 955 ATAGAGACTCGGGTTATCCGAGCCCGCATCTACTTGAATTTCTAGAACAACTTACAAAT 1014
Db 1027 ATAGAGGCTGCCATTTTTCAGGCTCCGCTATCTACTTGAATTTTCCAGAACAACTTACAAAT 1086
Qy 1015 TTTAGCACTTCATACGATGGAGTGTCTACTAGGCAATGACTTTACTGGCGGGGGCACACA 1074
Db 1087 TACAGTGCAATCAAGCGGTGGAGTAGCACTCAACATATGAATTAATTTGGGTGGGACATAGG 1146
Qy 1075 ATTCAATCTCGGCCAATAGGAGGCGGATTAATACCTCAACGATGGGTCTACCA---AT 1131
Db 1147 CTTAACTTCCGCCCATAATAGGAGGACATTAATACTCTCAACACAGGACTTACTAATTAAT 1206

Qy 1132 ACTTCTATTAATCCTGTAAGATTATCAATTTCTCTCGAGACGTATATTTGGACTGAATCA 1191
Db 1207 ACTTCAATTAATCCTGTAAGATTATCAATTTCTCTCGAGACGTATATTTAGAACAGATCA 1266
Qy 1192 TATGAGAGGTGCTTTCTATGCGGAATTTACCTTGAAACCTATTCTATGGTGTCCCTACTGTT 1251
Db 1267 AATGAGGAGCAAAATAT-----ACTATTTACTACTCTCTGTAATGGAGTACTTTGGGCT 1320
Qy 1252 AGATTTAATTTTATGGAACCCCTCAGAACTATTTTGAAGAGGTACTGCTAACTATAGTCAA 1311
Db 1321 AGATTTAATTTTATGGAACCCCTCAGAACTATTTTGAAGAGGCGCCACTACTCTACAGTCAA 1380
Qy 1312 CCCTATAGTCACTCTGGGCTTCAATTTAAAGATTCAGAAATCTGAATTAACCAACAGAAACA 1371
Db 1381 CCGTATCAGGAGTGGGATTTCAATTTATTTGATTCAGAACTGAATTAACCAACAGAAACA 1440
Qy 1372 ACAGAACGACCAAAATTTAGATCATATAGTCAATAGGTTATCTCATATAGGCTCATTTCA 1431
Db 1441 ACAGAACGACCAAAATTTAGATCATATAGTCAATAGTCAATAGTCAATATAGGACTAATCAT 1500
Qy 1432 CAATCTAGGCTGCATGTACCAAGTATTTCTTTGGAGCGCACCTGTAGTGCAGATCGTCAAT 1491
Db 1501 GGAACACTTTTGAGAGCACCAAGTCTATTTCTTTGGAGCGCATCGTAGTGCAGATCGTCAAT 1560
Qy 1492 ACCATTAGTTCAGATAGCATAAACAAATACCATTTGGTAAATCATCTCAACCTTAATTTCA 1551
Db 1561 ACGATTGCAACCAATAGATTTACAAATACATCTGTTGTTAAAGCACTGATTTCTTCAATCA 1620
Qy 1552 GGTACTCTGTAGTCAAGTGCAGGATTTTACAGAGGGGATATAATCCGAACCTAACGTT 1611
Db 1621 GGTGTTACTGTTGTTGGAGGCGCAGGATTTACAGTGGGGATATCCTTCGTAGAACAAAT 1680
Qy 1612 AATGGTAGTGTACTAAGTATGGTCTTAATTTTAAATAATCATCATCTACAGCGGTATCGC 1671
Db 1681 ACGGTCATTTGGAGATATACGATTAATTAATTAATGTGCCATTTATCCAAAGATATCGC 1740
Qy 1672 GTGAGAGTTCGTTATGCTTCTCAACAAATGGTCTGAGGGTAACTGTGCGAGGGAGT 1731
Db 1741 GTAGGATTCGTTATGCTTCTACTACAGATTTACAAATTTTTCACGAGATTAATGGAACC 1800
Qy 1732 ACTACTTTTGTATCAAGGATTCCTTAGTACTATAGTGCAAATAGTCTTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATGGTAAATTTCTCAAGAACTATGAATAGGGGGGATAATTTAGAAATATAGA 1860
Qy 1792 TCATTTAGATTTGAGAAATTTCTGTTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
Db 1861 AGTTTGAAGCTGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCNAAGCACTTC 1920
Qy 1852 ATAAAGTATAAGTAAATATGACAGGTAGACAAACGTTTCACTTTGATATAAATTTGAATTCAT 1911
Db 1921 ACATTTGGGTGCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGAGTTCGAATTTGTT 1980
Qy 1912 CCAATTAAGTGCACCTTCGAAAGCAGAAATACGATTTAGAAAGGGCGCAAGAGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAAACATTTGAGGCGAAGATATGATTTAGAAAGAGCACAAGAGCGGTGAAT 2040
Qy 1972 GCTCTGTTTACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTTATCATATT 2031
Db 2041 GCTCTGTTTACTTCTTACAAATCCAAAGAGATTGAAACAGATGTGACAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTATCGGATGAATTTCTGCTTAGATGAAAGAGA 2091
Db 2101 GACCAAGTGTCCAATATGGTGGCATGTTTATCAGATGAATTTTGTCTGATGAGAGGCA 2160
Qy 2092 GAATTTACTTGAGAAAGTGAATATTCGAAACGACTCAGTGTATGAAAGAACTTACTCCAA 2151
Db 2161 GAAATTTTGGAGAAAGTGAATATTCGAAAGCGACTCAGTGTATGAAAGAACTTACTCCAA 2220
Qy 2152 GATCCAAACTTCAATCCATCAATTAAGCAACGACTTCAATCTACTAATGAGCAATCG 2211
Db 2221 GATCCAAACTTCAATCCATCAATTAAGTGTGGGCAATTAAGTTTTCGATCCATCGATGGGCAATCA 2280

CC midgut cells, resulting in cessation of feeding, and eventually death.
CC The CryI proteins produced by B.t. are active against lepidopteran
CC insects. The protein encoded by this sequence, and the CryE4 protein
CC (see AAW17700) belong to the CryI family of ICPs. This sequence can be
CC used to transform bacteria, which are useful as insecticides against a
CC wide range of lepidopteran pests, and can be applied to crops, soil and
CC seeds. The encoded protein, or especially its toxic N terminal region,
CC can be expressed in plants, to provide protection against lepidopteran
CC pests. This sequence, or its fragments, can also be used to isolate other
CC similar genes. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ

Sequence 3934 BP; 1347 A; 650 C; 841 G; 1096 T; 0 U; 0 Other;

Query Match 72.9%; Score 2687.2; DB 2; Length 3934;

Best Local Similarity 83.5%; Pred. No. 0;

Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

Qy	1	TTGACTTCAATAGAGAAATAGAAATGAAATATATAATATGCTTTATCGATTCACGCTGA	60
Db	67	TTGACTTCAATAGAGAAATAGAAATGAAATATATAATGCTTTATCGATTCACGCTGA	126
Qy	61	TGCAATCATCCACAAATGATCTATCACCAGATGCTCGTATTGAGGATTCCTTGTGT	120
Db	127	TGCAATCCTTCCACCAATGATCTATCACCAGATGCTCGTATTGAGGATTCCTTGTGT	186
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	187	GTAGCCGAGGTGAACAAATATTGATCCATTTGTAGCGCATCAACAGTCCAAACGGGTATA	246
Qy	181	AACATAGCTGGTAGAATAGTGTATAGGCGTACCGTTTGTGGCAAAATAGCTAGT	240
Db	247	AACATAGCTGGTAGAATATGGGCGTATATAGGTTGGCGTTTGTGGCAAACTAGCTAGT	306
Qy	241	TTTTATAGTTTCTTGTGGTGAATATATGGCCCGCGGAGAGATCACTGGGAAATTTTC	300
Db	307	TTTTATAGTTTCTTGTGGGGAATATATGGCCCGTAGTGGCAGAGATCCATGGGAAATTTTC	366
Qy	301	CTAGAACTGTGCAACAACTTATAAATCAACAAATAACAGAAATGCTTAGGAATACGGCA	360
Db	367	CTAGAACTGTAGAACAACTTATAGACAAACAGTAAACAGAAATGCTTAGGAATACGGCT	426
Qy	361	CTTGCTCGATTAACAGGTTTAGGAGATTCCTTTTAGAGCTTATCAACAGTCACTGAAGAT	420
Db	427	ATTGCTCGATTAACAGGTTTAGGAGATTCCTTTTAGAGCTTATCAACAGTCACTGAAGAT	486
Qy	421	TGGCTAGAAACCGTGATGATGCAAGACGAGAGTGTCTTTATACCAATATATAGCC	480
Db	487	TGGTTAGATAACCGAAATGATGCAAGATCAAGAGCATTTCTTTGAGCGCTATGTTGCT	546
Qy	481	TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	540
Db	547	TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	606
Qy	541	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTTATGAGAGATGCTCT	600
Db	607	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTTATGAGAGATGCTCT	666
Qy	601	CTTTTGGTGTAGTATTTGGGCTTACATCGCAGGAAATCAACGTTATTATGAGGCGCAA	660
Db	667	CTTTTGGTGTAGTATTTGGGCTTACATCGCAGGAAATCAACGTTATTATGAGGCGCAA	726
Qy	661	GTGGAACAAACGAGAGATTTATTCGAGCTATTGCGTAGAATGGTATAATACAGGTTCTAAT	720
Db	727	ATCAGATATACAGAGAAATTTCTAACCATTTGCGTAGAATGGTATAATACAGGTTCTAAT	786
Qy	721	AGCTTTAGAGGACAAATGCGCGAAGTTGGGTGGTATATCAATTCCTCGTAGAGATCTA	780
Db	787	AACTTAGAGGACAAATGCTGAAAGTTGGTTGGGTATATCAATTCCTCGTAGAGATCTA	846
Qy	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA	840
Db	847	ACGTTAGGGGTATTAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA	906

Qy	841	ATAAATACAGTGTCTCAGTTAAACAGGAAGTTTATACAGACGCAATTTGGAGCAACAGG	900
Db	907	ATCATACAGTGTCTCAGTTAAACAGGAAGTTTATACAGATCCAAATTTGGGAGCAAAAT	966
Qy	901	GTAAAT-----ATGGCAAGTATGAATGGTATATAATATAATATGCACTTCGTTTCCGCT	954
Db	967	GCACCTTCAGGATTTGCAAGTACGAATGGTTTAAATAATAATATGCACTTCGTTTCCGCT	1026
Qy	955	ATAGAGCTGGGTATTCGGAAGCCGATCTACTGATTTTCTAGAACAACTTCAATTT	1014
Db	1027	ATAGAGCTGGCATTTTTCAGGCTCCGATCTACTTGAATTTTCCAGAACAACTTCAATTT	1086
Qy	1015	TTTAGCACTTCATCAGATGAGTGTCTACTAGGCATATGACTTACTGGGGGGGACACACA	1074
Db	1087	TACAGTGCATCAAGCCGTTGGAGTAGCACTCAACATATGAATTTATGGTGGGACATAGG	1146
Qy	1075	ATTCAATCTCGGCAATAGGAGGGGATTAATACCTCAACGATGGGTCTACCA---AT	1131
Db	1147	CTTAACCTTCGCCCAATAGGAGGGGACATTAATAATACCTCAACAGAGGACTTACTAATAAT	1206
Qy	1132	ACTTCTATTAATCTGTGAAGATTATCACTTCTCTCGAGAGCTATATTGGACTGAATCA	1191
Db	1207	ACTTCAATTAATCTGTAAACATTACGTTTACGCTCTCGAGAGCTTTATAGAACAGAACTCA	1266
Qy	1192	TATGAGAGTGTCTTATGGGGAAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTT	1251
Db	1267	ATGAGGGGCAAAATAT-----ACTATTACTACTCTGTGAATGGAGTACTTGGGCT	1320
Qy	1252	AGATTTAATTTTGAAGACCCCTCAGAATACTTTTGAAGAGGTACTGTCTAATATAGTCAA	1311
Db	1321	AGATTTAATTTTAAACCCCTCAGAATACTTTTGAAGAGGGCGGCACTTACTACAGTCAA	1380
Qy	1312	CCCTATGAGTCACTGGGCTTCAATTAAGAACTTCAAGAACTGAATTTACCAACAGAAACA	1371
Db	1381	CCGTATCAGGAGTGGGATTTCAATTTATTTGATTCAGAACTGAATTTACCAACAGAAACA	1440
Qy	1372	ACAGAACGACCAAAATTAATATATAGTATATAGTGTCTCATAGGCTCTAATTTCA	1431
Db	1441	ACAGAACGACCAAAATTAATATATAGTATATAGTGTCTCATATAGGCTCTAATTTCA	1500
Qy	1432	CAATCTAGGTTGATGATCCAGTATATTTCTTGGAGCGACCGTAGTGCAGATCGTACAAAT	1491
Db	1501	GGAAACACTTTTGAGAGCACCACTCTATTTCTTGGAGCGATCGTAGTGCAGATCGTACGAAT	1560
Qy	1492	ACCAATGTTGATGATTAACCAATACCAATACCAATTTGTAATCTTCAACCTTAATTTCA	1551
Db	1561	ACGAATGGACCAAAATAGAAATTAACAAATACCAATTTGTAATTTGTAATTTCAATTTCA	1620
Qy	1552	GGTACCTCTGTAGTGCAGTGGCCAGGATTTTACAGAGGGGATATAATCCGAACTAAACGTT	1611
Db	1621	GGTGTACTGTGTGTGGAGGGCAGGATTTTACAGTGGGATATCTCTGTAGAACAAAT	1680
Qy	1612	AATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1671
Db	1681	ACGGGTACATTTGAGAGATACGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT	1740
Qy	1672	GTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1731
Db	1741	GTAAGGATTCGTTATGCTTCTACTACAGATTTTCAATTTTTCACGAGAAATTAATGGAAC	1800
Qy	1732	ACTACTTTTGTATCAAGGATTTCCCTAGTATGATGATGATGATGATGATGATGATGATGATGAT	1791
Db	1801	ACTGTTAATTTGTTAATTTCTCAAGAACTATGATAGGAGGGGATTAATTTAGNATATAGA	1860
Qy	1792	TCATTTAGATTTGAGAAATTTCTGTAGTATAGTGCATCTGGCAGTCAAACTGCTGGA	1851
Db	1861	AGTTTAAAGACGAGGATTTAGTACTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT	1920
Qy	1852	ATAAGTATAGTATATAGTGCAGGTTAGACAAACGTTTCACTTTGATTAATTTGAATTTCAAT	1911
Db	1921	ACATTTGGGTGCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGAGTGCAGATTTGTT	1980
Qy	1912	CCAATTTACTGCAACCTTCTGAGGAGAAATACGATTTTAGAAGGGCGGCAAGGCGGTGAAT	1971

[illegible]

```
Db 1741 GTAAGGATTCGTTATGCTTCTACTACAGATTACAATTTTTCACGAGAATTAATGAACC 1800
Qy 1732 ACTACTTTTGAATCAAGAGATTCCTAGTACTATGAGTGCAAAATGAGTCTTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATGTTAATTTCTCAAGAACTATGAATAGGGGGGATAATTTTGAATAATAGA 1860
Qy 1792 TCATTTAGATTTCAGNAATTTCTGTAGGTATTTAGTGCACTCTGGCAGTCAAACGTCTGGA 1851
Db 1861 AGTTTGTAGAACTCGAGATTAGTACTCTCTTTTAATTTTAAATGCCCAAGCACATTC 1920
Qy 1852 ATAAGTATAAGTAATAATGAGGTAGACAAACGTTTCACATTTTCATAAAATTTGAATTCATT 1911
Db 1921 ACATGGGTCTCAGAGTTTTCAATCAGGAAGTTTATATAGATAGATCGAATTTGTT 1980
Qy 1912 CCAATTAATCAGACCTTTGAGAGAGAAATACGATTTTAAAGAGCGCAAGAGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAAACATTTGAGGCAGAAATATGATTTAGAAAGACACAAAGGCGGTGAAT 2040
Qy 1972 GCTCTGTTTAACTAATACGAATCCAAGNAGATTGAAAACAGATGTGACAGATTAATCATATT 2031
Db 2041 GCTCTGTTTAACTTCTCAAAATCCAAGAAAGATTGAAAACAGATGTGACAGATTAATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGCGGTGTTATCGGATGAATTTCTGCTTAGATGAAAAGAGA 2091
Db 2101 GACCAAGTGTCCAATATGTTGGCATGTTTATCAGATGAATTTTGTCTTGGATGAGAAGCGA 2160
Qy 2092 GAATTAATCTGAGAAAGTGAATATGCGAAACGACTCAGTGTGATGAAAGAACTTACTCCAA 2151
Db 2161 GAATTAATCTGAGAAAGTGAATATGCGAAGCGACTCAGTGTGATGAAAGAACTTACTCCAA 2220
Qy 2152 GATCCAAACTTCACTCAATCAATAGCACACAGACTTCATATCTACTAATGAGCAATCG 2211
Db 2221 GATCCAAACTTCACTTCACTTCACTGAGCAATTAAGTTTTCGCATCCATCGATGGCAATCA 2280
Qy 2212 AATTTTCACTATCTCATGAACCAATCTGAAACATGGATGTTGGGGAAGTGAGAACATTACA 2271
Db 2281 AACTTCCCTCTAATTAATGAGCTATCTGAACATGGATGTTGGGGAAGTGCGAATGTTACC 2340
Qy 2272 ATCCAGAAAGAAATGACGTATTTTAAAGAGAAATACGTACACACTACCGGGGACTTTTAAT 2331
Db 2341 ATTCCAGAAAGGAATGACGTATTTTAAAGAGAAATACGTACACACTACCGGGTACTTTTAAT 2400
Qy 2332 GAGTGTTATCCGACGTATTTATATCAAAATAGGAGAGTCGGAATTTAAAGCTTACT 2391
Db 2401 GAGTGTTATCCAAATTAATTTATATCAAAATAGGAGAGTCAGAAATTTAAAGCTTATACG 2460
Qy 2392 CGCTACCAATTTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGT 2451
Db 2461 CGCTATCAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATTTATTTAATTCGT 2520
Qy 2452 TATAATCGCAAAATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGTATGGCCGCTTCA 2511
Db 2521 TACAATGCAAAAGCATGAAACATTTGGATGTTTCCAGGTACCGATTTCCCTATGGCCGCTTCA 2580
Qy 2512 GTTGAAGCCCAATCCGAGGTCGGAGACCGAATCGATGCGCACACATTTTGAATGG 2571
Db 2581 GTTGAAGCCCAATCCGAGGTCGGAGAACCAAAATCGATGCGCACACATTTTGAATGG 2640
Qy 2572 AATCCTGATCTAGATTGTTTCCCTCAGAGATGGAGAAAATTTGCGCATCATTTCCCATCAT 2631
Db 2641 AATCCTGATCTAGATTGTTTCCCTCAGAGATGGAGAAAGATGTGCGCATCATTTCCCATCAT 2700
Qy 2632 TTCTCTTTGGATATTTGATATTTGATGACAGACTTGTGATGAGAAATCTTAGCGGTGGGTG 2691
Db 2701 TTCACTTTGGATATTTGATGTTGGGTGCACAGACTTGCATGAGAACTTAGCGGTGGGTG 2760
Qy 2692 GTATTCAAGATTAGACGCAAGGAGTTCATGCAAGACTAGGGAATCTGGAATTTATTGAA 2751
Db 2761 GTATTCAAGATTAGACGCAAGGAGGTTATGCAAGATTAGGAATCTGGAATTTATCGAA 2820
Qy 2752 GAGAAACCATTTATAGGAGAAAGCACTGCTCTGTTGAAAGAGACAGAGAAAAAATGGAA 2811
```

```
Db 2821 GAGAAACCATTTAATTTGGAGAACGACGTCTCGTGTGAAGAGACGCGAABAAAATCGAGA 2880
Qy 2812 GACAAACGTGAAAAAACAATTTGGAACAAACACGAGTATATACAGAGCGCAAAAGAGCT 2871
Db 2881 GACAAACGTGAAAAAACAATTTGGAACAAACACGAGTATATACAGAGCGCAAAAGAGCT 2940
Qy 2872 GTGGATGCTTTTATTTGTAGATTTCTCAATATATAGATTTACAAGCGGATACAAACATTGGC 2931
Db 2941 GTGGATGCTTTTATTTGTAGATTTCTCAATATGATCAATTTACAAGCGGATACAAACATTGGC 3000
Qy 2932 ATGATTTCAATCGCGCAGATAAACTTGTTCATCGAATTTCCGAGAGCTTATCTGTGAGAAATTA 2991
Db 3001 ATGATTTCAATCGCGCAGATAAACTTGTTCATCGAATTTCCGAGAGCGTATCTTTCAGAAATTA 3060
Qy 2992 TCTGTTATCCCGGGTGTAAATCGGGAATTTTGTGAAGAAATTAAGAGGTGCGATTATCACT 3051
Db 3061 CCTGTTATCCAGGTGTAAATCGGGAATTTTGTGAAGAAATTAAGAGGTGCGATTATCACT 3120
Qy 3052 GCAATCTCCTATACGATCGAGAAATGTCGTTAAAAATGGTGATTTTAAATATGATTA 3111
Db 3121 GCAATGTCCTTATACGATCGGAAATGTCGTTAAAAATGGTGATTTTAAATATGATTA 3180
Qy 3112 GCATGCTGGAATGTAAAAGGCGATGTAGATGTACAAACAGAGCCATCACCGTTCTGTCCTT 3171
Db 3181 ACATGTTGGAATGTAAAAGGCGATGTAGATGTACAAACAGAGCCATCATCGTTCTGACCTT 3240
Qy 3172 GTTATCCAGAAATGGGAGCAGAAAGTGTCAAGACGATTCGCGTCTGTCGGGGCGGTGC 3231
Db 3241 GTTATCCAGAAATGGGAGCAGAAAGTGTCAAGACGATTCGCGTCTGTCGGGGCGGTGC 3300
Qy 3232 TATATCTCCGTGTCAACAGCGTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCAT 3291
Db 3301 TATATCTTCTGTCACAGCGTACAAAGAGGATATGGAGAGGTCGCGTAACGATCCAT 3360
Qy 3292 GAAATCGAACAATACAGACGAACTAAAATTTAAAAAATCTGCAAGAGAGGAAAGTGTAT 3351
Db 3361 GAAATCGAACAATACAGACGAACTAAAATTTAAAAAATCTGCAAGAGAGGAAAGTGTAT 3420
Qy 3352 CCACCGGATACAGGAACGTTAATGATTAATCTGCACACCAAGGTACAGCATGTGTAAT 3411
Db 3421 CCACCGGATACAGGAACGTTAATGATTAATCTGCACACCAAGGTACAGCATGTGTAAT 3480
Qy 3412 TCCGTTAATGCTGGATATGAGGATGCATATGAAGTTTGATCTACAGCATCTGTTAATTAC 3471
Db 3481 TCCGTTAATGCTGGATATGAGGATGCATATGAAGTTTGATCTACAGCATCTGTTAATTAC 3540
Qy 3472 AAACCGACTTATGAAGAAGAAAACGTATACAGATGTACGAAGAGATAATCATTTGTGAATAT 3531
Db 3541 AAACCGACTTATGAAGAAGAAAACGTATACAGATGTACGAAGAGATAATCATTTGTGAATAT 3600
Qy 3532 GACAGAGGTATGTGAATTTATCCACCATCCAGCTGGTTTATATGACAAAAGNAATTAGAA 3591
Db 3601 GACAGAGGTATGTGAATTTATCCACCATCCAGCTGGTTTATGTCAAAAGNAATTAGAA 3660
Qy 3592 TACTTCCCAAGAAACCGATAAGGTATGAGTTGAGATTGGAGAAACGGAAGGAGTTTATT 3651
Db 3661 TACTTCCCAAGAAACAGATACAGTATGGATTGAGATTGGAGAAACGGAAGGAGTTTATT 3720
Qy 3652 GTAGACAGCGTGGAAATTTACTCTTATGGAGGAATAG 3687
Db 3721 GTAGATAGCGTGGAACTACTCTCTCATGGAAGAAATAG 3756
```

RESULT 12

ADK98478
ID ADK98478 standard; DNA; 3687 BP.

XX ADK98478;

XX AC
XX 03-JUN-2004 (first entry)

XX DE B thuringiensis cryIIb partial gene sequence SeqID1.
XX

KW insecticidal protein; plant; pesticide; gene therapy;
 KW lepidopteran insect pest; transgenic plant;
 KW insect infestation resistance; monocot; dicot; cry1Bb; gene; ds.
 XX
 OS Bacillus thuringiensis.

Key Location/Qualifiers
 CDS 1..3687
 /tag= a
 /product= "B thuringiensis cry1Bb"
 /partial
 /note= "No start or stop codon"

WO2004020636-A1.

11-MAR-2004.

26-AUG-2003; 2003WO-US026510.

29-AUG-2002; 2002US-0407428P.

(MONS) MONSANTO TECHNOLOGY LLC.
 (BOGD/) BOGDANOVA N N.
 (ROMA/) ROMANO C P.

Bogdanova NN, Romano CP;

WPI; 2004-269221/25.

P-FSDB; ADK98479.

New polynucleotide sequence optimized for expression of an insecticidal protein in a plant, useful in the control of Lepidoptera insect pests, and for producing transgenic plants with the ability to resist insect infestations.

Disclosure; SEQ ID NO 1; 138pp; English.

This invention relates to a novel polynucleotide sequence optimised for expression of an insecticidal protein in a plant. The invention may be useful for the production of pesticides whilst the disclosed sequences may be used for gene therapy. The polynucleotide sequence and methods are useful in the control of lepidopteran insect pests, and for producing transgenic plants with the ability to resist insect infestations. The invention provides polynucleotide sequences with enhanced, improved and optimised expression in monocot and dicot plant species. The present sequence is that of the (partial) B thuringiensis cry1Bb gene which is related to the invention.

Sequence 3687 BP; 1248 A; 621 C; 796 G; 1022 T; 0 U; 0 Other;

Query Match 72.8%; Score 2684.2; DB 12; Length 3687;
 Best Local Similarity 83.5%; Pred. No. 0;
 Matches 3085; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

QY 1 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATATATAATGCTTTATCGATTCCAGCTGTA 60

DB 1 TTGACTTCAATAGGAAAAATGAGAAATGAAATATATAATGCTTTATCGATTCCACGTA 60

QY 61 TCGAATCAATCCACAATAATGAAATATATCCAGATGCTCGTATTGAGGATCTTTTGT 120

DB 61 TCGAATCTTCCACGAAATGAATCTATCAACAGATGCTCGTATTGAAAGATAGCTTGT 120

QY 121 ATAGCCGAGGGGAATAATATCAATCACTTGTAGCGCATCAACAGTCCAAACGGGTATT 180

DB 121 GTAGCCGAGGGTGAACAAATATTGATCCATTGTGTAGCGCATCAACAGTCCAAACGGGTATA 180

QY 181 AACATAGCTGGTAGAATACTAGTGTATTAGCGGTACCGTTTGTCTGGACAAATAGCTAGT 240

DB 181 AACATAGCTGGTAGAATAATTGGCGGTATTAGTGTGCCGTTTGTCTGGACAACTAGCTAGT 240

QY 241 TTTTATAGTTTTCTTGTGGTGAATATTAGCCCGCGGCGAGAGATCAGTGGGAAATTTTC 300

DB 241 TTTTATAGTTTTCTTGTGGGGAATATTAGCCCTAGTGGCAGAGATCCATGGGAAATTTTC 300

QY 301 CTAGAAATGTCGAAACAACTTTATAAATCAACAAATAACAGAAAAATGCTAGGAATACGGCA 360
 DB 301 CTGGAACATGTAGAAACAACTTTATAAGACACAACTAACAGAAAAATCTAGGAATACGGCT 360
 QY 361 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTTAGAGCTTATCAACAGTCACTTTGAAGAT 420
 DB 361 ATTGCTCGATTAGAAGGTTTAGGAAAGAGGCTTATAGATCTTTACCAGAGGCTCTTTGAAACT 420
 QY 421 TGGCTAGAAACCGTGTATGATGCAAGAACGAGAGAGTGTCTTTATACCAATATATATAGCC 480
 DB 421 TGGTTAGATTAACCGAAATGATGCAAGATCAAGAGCATTTCTTTGAGCGCTATGTTGCT 480
 QY 481 TTAGAACTTGATTTTCTTAATGCCATGCCGCTTTTCGCAATTAGAAAAACAAAGAGTTTCCA 540
 DB 481 TTAGAACTTGATTTTCTTAATGCCATGCCGCTTTTCGCAATTAGAAAAACAAAGAGTTTCCA 540
 QY 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTTACCTTATTTATTATTAGAGATGCTCTCT 600
 DB 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTTACCTTATTTATTATTAGAGAGCGCATCC 600
 QY 601 CTTTTTGTAGTGAATTTTGGGCTTACATCGCAGGAAATTTCAACGTTATTATTAGCGCCAA 660
 DB 601 CTTTTTGTAGTGAATTTTGGGCTTACATCGCAGGAAATTTCAACGTTATTATTAGCGCCAA 660
 QY 661 GTGGAACAAACGAGAGATTTATTCGACTATTTCGCTAGATGCTATATATACAGGCTCTAAAT 720
 DB 661 ATCAGATATACAGAGGAAATTTCTAACATTTGCTGCAATGCTATATATACAGGCTCTAAAT 720
 QY 721 AGCTTGAGAGGACAAATGCGCAAGTTGGGTGCTTTATTAATCAATTCGCTAGAGATCTA 780
 DB 721 AACTTAAGAGGACAAATGCTGAAAGTTGGTTGGGTATTAATCAATTCGCTAGAGACCTA 780
 QY 781 ACGTAGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACATCGGCACATTATCCA 840
 DB 781 ACGTAGGGTATTAGATTTAGTAGCCCTATTCCCAAGCTATGATCTCGCACCTTATCCA 840
 QY 841 ATAAATACGAGTGTCTAGTTTAAACAGGAAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
 DB 841 ATCAATACGAGTGTCTAGTTTAAACAGGAAAGTTTATACAGATCCCAATTTGGAGCAACAAAT 900
 QY 901 GTAAAT-----ATGCAAGTATGAATTTGGTATTAATTAATATGCACTTCGTTTTCGCT 954
 DB 901 GCACCTTCAGGATTTGCAAGTATGCAATTTGGTTTAAATAATATGCACTTCGTTTTCGCT 960
 QY 955 ATAGAGACTGGGTTATCCGAGCCGCTACTGTTGTTTCTTAGAACAACTTACAAAT 1014
 DB 961 ATAGAGCTGCCATTTTCAGGCTCCGCTACTGTTGTTTCCAGAACAACTTACAAAT 1020
 QY 1015 TTTAGCACTTCATCAAGATGAGTGCTACTAGGCTATGACTTACTGCGGGGGCACACA 1074
 DB 1021 TACAGTGCATCAAGCCGTTGGAGTAGCACTCAACATATGAATTTATGGTGGGACATAGG 1080
 QY 1075 ATCAATCTCGGCCAATAGGAGCGGATTAATACTCAACGCAATGG---GTCTACCAAT 1131
 DB 1081 CTTAACTTCGCCCCAATAGGAGGACATTAATACTCAACACAAAGGACTTACTTAATAAT 1140
 QY 1132 ACTTCTAATTAATCTGTAAAGATTAATCACTTCTCTCGAGACGATATATTGAGACTGAATCA 1191
 DB 1141 ACTTCAATTAATCTGTAAATTAACAGTTTACGCTCTCGAGACGTTTATAGAACAGAAATCA 1200
 QY 1192 TATCAGAGGCTTCTATGCGGGAATTTTACCTTGAACCTATTTCATGTTGTCCTCTACTGTT 1251
 DB 1201 AATCAGGACAAATAT-----ACTATTTACTCTCTGTAATGGAGTACCTTTGGGCT 1254
 QY 1252 AGATTTAATTTTAGGAACCTCTCAGAATCTTTTGAAGAGGACTACTGTCTAACTATAGTCAA 1311
 DB 1255 AGATTTAATTTTAAACCTCTCAGAATATTATTAAGAGAGCGCCACTTACCTACAGTCAA 1314
 QY 1312 CCCTATGAGTCACTGGGCTTCAATTAAGAGTTTCAAGAACTGAATTAACCAACAGAAACA 1371
 DB 1315 CCGTATCAGGAGTTGGGATTTCAATTTATTTGATTCAGAAACTGAATTAACCAACAGAAACA 1374

1372 ACAGAACGACAAATTATGAATCATATAGTCATAGTTTCTCACATAGGCTCATTTCA 1431
1375 ACAGAACGACAAATTATGAATCATATAGTCATAGTTTCTCATATAGGACTTAATCATATA 1434
1432 CAATCTAGGGTGCAATGACAGATATATCTTGGACGACCGTAGTCAGATCGTACAAAT 1491
1435 GGAAACACTTTGAGAGCACCAGTCTATTCTTGGACGCATCGTAGTCAGATCGTACGAAT 1494
1492 ACCATTAGTTCAGATAGCATAAACAAATACCAATGGTAAATCAATCAACCTTAATTC 1551
1495 ACATTGGACCAATAGAAATACAAATACCAATGGTAAAGACCTGAATCTTCAATTC 1554
1552 GGTACCTCTGTAGTCAGTGCCGACAGATTTACAGGAGGGGATATAATCGAACTAACGTT 1611
1555 GGTGTACTTGTGTGGAGGCGCAGATTTACAGGTGGGGATATCCTTGTAGAACAAAT 1614
1612 AATGGTAGTGACTAAAGTAGTGGTCTTAATTTTAAATAATCAATCATACAGCGGTATCG 1671
1615 ACGGCTACATTTGGAGATATACGATTAAATATTAATGTGCCATTAATCCCAAGATATCG 1674
1672 GTCAGAGTTCGTTATGCTGCTCTCAACAAATCGTCTGAGGTAATGTCGAGGGAGT 1731
1675 GTAAGGATTCGTTATGCTCTACTACAGATTTTCAATTTTTCACGAGAAATTAATGGAAC 1734
1732 ACTACTTTTGGATCAAGGATTTCCCTAGTACTATGAGTGCAATGAGTCTTTGACATCTCAA 1791
1735 ACTGTTAATATGGTAATTTCTCAAGAACTATGATAGGGGGATTAATTTAGATATAGA 1794
1792 TCATTTAGATTTCCAGAAATTTCTGTAGGTATTTAGTGCACTCTGGCAGTCAAACTGCTGGA 1851
1795 AGTTTAAAGACTCGAGATTTAGTACTCTCTTTTAAATTTTAAATGCCCAAGCACATTC 1854
1852 ATAAGTATAGTAATTAATGAGGTAGACAAACGTTTTCATTTGATATAAATTTGAATTCAT 1911
1855 ACATTTGGTGCTCAGAGTTTTCCTCAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1914
1912 CCAATTAATGCAACCTTTGAGGAGAGATACGATTTAGAAAGGCGCAAGGCGGTGAAT 1971
1915 CCAGCAGAGGTAACTTTGAGGCGAGATATGATTTAGAAAGAGCAACAAGGCGGTGAAT 1974
1972 GCTCTGTTTACTAATACGAATCCAAAGATTTGAAACAGATGTGACAGATATCATATTT 2031
1975 GCTCTGTTTACTTCTCAATCCAAAGATTTGAAACAGATGTGACAGATATCATATTT 2034
2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTCGATGGAATTCGTTAGATGAAGAGAG 2091
2035 GACCAAGTGTCCAAATATGTTGGCATGTTTATCAGATGAATTTTGTGATGAGAGCGA 2094
2092 GAATTAATTTGAGAAAGTGAAATATGCGAAACGACTCAGTGATGAAGAAAGACTTACTCCAA 2151
2095 GAATTAATTTGAGAAAGTGAAATATGCGAAGCGACTCAGTGATGAAGAAAGACTTACTCCAA 2154
2152 GATCCAAACTTACATCCATCAATAGCAACAGACTTCATATCTAATAGAGCAATCG 2211
2155 GATCCAAACTTACATCCATCAGTGGGCAATTAAGTTTCGATCCATCGATGCAACAATCA 2214
2212 AATTTCCACATCTATCCATGAACAACTCTGAACATGGATGTTGGGGAAGTGAGACATTACA 2271
2215 AACTTCCCCTCTAATAGAGCTATCTGAACATGGATGTTGGGGAAGTGCGAATGTTACC 2274
2272 ATCCAGGAAGGAATGACGCTATTTAAAGAGAAATTAAGTCACTACCGGGGACTTTTAAT 2331
2275 ATTCAGGAAGGAATGACGCTATTTAAAGAGAAATTAAGTCACTACCGGGTACTTTTAAT 2334
2332 GAGTGTTATCCGAGCTATTTATATCAAAATAAGGAGAGTCCGAATTAAGAGCTTATCT 2391
2335 GAGTGTTATCCAAATTAATTTATATCAAAATAAGGAGAGTCCGAATTAAGAGCTTATACG 2394
2392 CGCTTCAAAATTAAGAGGATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGT 2451
2395 CGCTATCAATTAAGAGGATATTTGAAGATAGTCAAGATCTAGAGATTTATTTAATTCGT 2454
2452 TATAATGCGAACAATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATATGGCGCGTTTCA 2511

2455 TACAATGCAAGACATGAACATTTGGATGTTCCAGGTACCGATTCCTATGGCGCGCTTCA 2514
2512 GTTGAAGCCCAATCGGAAGGTGCGAGAAACCGAATTCGATGCGCACCAATTTTGAATGG 2571
2515 GTTGAAGCCCAATCGGAAGGTGCGAGAAACCAATTCGATGCGCACCAATTTTGAATGG 2574
2572 AATCTGATCTAGATTTGTTCTTCCGAGAGATGAGAAATAATGTCGCGATCATTTCCCATCAT 2631
2575 AATCTGATCTAGATTTGTTCTTCCGAGAGATGAGAAATAATGTCGCGATCATTTCCCATCAT 2634
2632 TTCTCTTTTGGATTTGATATTTGGATGTCAGACTTTGTCATGAGAATCTAGCGTGTGGGTG 2691
2635 TTCACTTTTGGATTTGATGTTGGGTGTCAGACTTTGTCATGAGAACCTAGCGGTGTGGGTG 2694
2692 GTATTCAAGATTTAAGACGCAAGGTCATGCAAGCTAGGGAATCTGGAATTTATTGAA 2751
2695 GTATTCAAGATTTAAGACGCAAGGTTATGCAAGATTTAGGAATCTGGAATTTTATCGAA 2754
2752 GAGAACCAATTTATTAGGAGAGCACTGCTCTGTTGAGAGAGACGAGAAAAAATGAGAA 2811
2755 GAGAACCAATTTATTGGAGAGCACTGCTCTGTTGAGAGAGACGAGAAAAAATGAGAA 2814
2812 GACAAACGTTGAAAACTACAATTTGGAACAAACCGAGTATATACAGAGGCAAAAAAGAGCT 2871
2815 GACAAACGTTGAAAACTACAATTTGGAACAAACCGAGTATATACAGAGGCAAAAAAGAGCT 2874
2872 GTGGATGCTTTTATTTGTAGATTTCTCAATATAATAGATTAACAAGCGGATACAAACATTGGC 2931
2875 GTGGATGCTTTTATTTGTAGATTTCTCAATATGATCAATTTACAAGCGGATACAAACATTGGC 2934
2932 ATGATTCATGCGCGAGATAAATCTTCTCATGAAATTCGAGAGGCTTATCTGTCAGAAATTA 2991
2935 ATGATTCATGCGCGAGATAAATCTTGTTCATGAAATTCGAGAGGCGTATCTTTCAGAAATTA 2994
2992 TCTGTTATCTCCCGGTGTAATCGGAAATTTTGAAGAAATTAAGAAGTTCGCAATTTATCACT 3051
2995 CCGTTATCTCCAGGTGTAATCGGAAATTTTGAAGAAATTAAGAAGTTCACATTTATCACT 3054
3052 GCAATCTCCATATACATGCGAGAAATGTCGTTAAAAATGGTGATTTTAAATAGGATTA 3111
3055 GCAATGTCCTTATACATGCGAGAAATGTCGTTAAAAATGGTGATTTTAAATAGGATTA 3114
3112 GCATGCTGGAATGTAAGAGGCGATGTAGATGTACAACAGAGCCATCACCGTCTCTGCTT 3171
3115 ACATGTTGGAATGTAAGAGGCGATGTAGATGTACAACAGAGCCATCATCGTTCTGACCTT 3174
3172 GTTATCCCAAGATGGAAGCAGAAAGTGTCACAAGCAGTTTCGCTCTGTCCGGGCGGTGCG 3231
3175 GTTATCCCAAGATGGAAGCAGAAAGTGTCACAAGCAGTTTCGCTCTGTCCGGGCGGTGCG 3234
3232 TATATCTCCGCTGTACAGCGTACAAGAGGATATGAGAGGTTGTGTAGGATCCAT 3291
3235 TATATCTCTGCTGTACAGCGTACAAGAGGATATGAGAGGCGTGTGTAACGATCCAT 3294
3292 GAAATCGAACAATACAGACGAACTAAAAATTTAAAACTGTGAAGAGAGGAAGTGAT 3351
3295 GAAATCGAACAATACAGACGAACTAAAAATTTAAAACTGTGAAGAGAGGAAGTGAT 3354
3352 CCAACGGATACAGGAACGTTGATTAATGATTACTGCACACCAAGGTACAGCAGTGTAAAT 3411
3355 CCAACGGATACAGGAACGTTGATTAATGATTACTGCACACCAAGGTACAGCAGTGTAAAT 3414
3412 TCCCGTAAATGCTGGATATGAGGATCGATATGAAGTTGATCTACAGCATCTGTTAATTAC 3471
3415 TCCCGTAAATGCTGGATATGAGGATCGATATGAAGTTGATCTACAGCATCTGTTAATTAC 3474
3472 AAACCGACTTATGAAGAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTGAATAT 3531
3475 AAACCGACTTATGAAGAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTGAATAT 3534
3532 GACAGAGGATGTTGAATTTATCCACCATACAGCTGCTGTTATATGACAAAAAGAAATTAGAA 3591

Db 3535 GACAGAGGCTATGTGAATTATCCACACAGTACCAGCTGGTATTGTGACAAAAGAAATTAGAA 3594
QY 3592 TACTTCCCAAGAACCCGATAGCTGATGAGATTGAGATTGGAGAAACGGAAGGGAGTTTATT 3651
Db 3595 TACTTCCCAAGAACACATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3654
QY 3652 GTAGACAGCGTGGAAATTACTCTTATGGAGGAA 3684
Db 3655 GTAGATAGCGTGGAACTACTCTCTATGGAAGAA 3687

RESULT 13
AAS04854
ID AAS04854 standard; DNA; 3627 BP.
XX
AC AAS04854;
XX
DT 07-SEP-2001 (first entry)
XX
DE Bacillus thuringiensis DNA encoding partial mutant Cry1Ba.
XX
KW Crystal protein; Cry1Ba; moth; butterfly; Colorado potato beetle;
KW ds; mutant.
XX
OS Bacillus thuringiensis.

XX Key Location/Qualifiers
PH 1. 3627
CDS /tag= a
FT /product= "Cry1Ba"
FT /partial
FT /note= "No stop codon"
FT sig_peptide 1. 57
FT /tag= b
FT mat_peptide 58. 3651
FT /tag= c
FT /label= Mature Cry1Ba
FT mutation replace(1464,A)
FT /tag= d
FT mutation replace(1467,T)
FT /tag= e

XX EP1099760-A1.
PN
XX
XX 16-MAY-2001.
XX
XX 09-NOV-1999; 99EP-00203723.
XX
XX 09-NOV-1999; 99EP-00203723.
XX (CPRO-) CPRO-DLO CENT PLANTENVERDEDELINGS REPROD.
XX
XX De Maagd RA, Bosch HJ;
XX
XX WPI; 2001-337141/36.
DR P-PSDB; AAU02094.

XX New hybrid Bacillus thuringiensis hybrid toxins comprising structural
PT domains derived from at least 2 different crystal proteins, such as
PT CryIIa and CryIbA, and having insecticidal activity, useful for combating
PT insects.
XX
XX Example; Page 22-23; 43pp; English.
XX
XX The sequence encodes B. thuringiensis (Bt) crystal protein CryIbA,
CC mutated to allow cloning of domain III or domains I and II, to make the
CC hybrid protoxins of the invention. The hybrid toxins of the invention,
CC having structural domains I, II and III in this order starting from the N
CC -terminal derived from at least 2 different crystal proteins, are useful
CC for protecting plants against pest insects, e.g. moths, butterflies and
CC Colorado potato beetle or for combating insects
XX
XX Sequence 3627 BP; 1184 A; 648 C; 820 G; 975 T; 0 U; 0 Other;
SQ

Query Match
Beat Local Similarity 71.3%; Score 2630.2; DB 4; Length 3627;
Matches 3081; Conservative 0; Mismatches 533; Indels 81; Gaps 5;
QY 2 TGACTTCAAAATAGGAAAAATGAGAAATTAATAAATGCTTTATCGAATTCAGCTGTAT 61
Db 2 TGACTTCAAAATAGGAAAAATGAGAAATTAATAAATGCTTTATCGAATTCAGCTGTAT 46
QY 62 CGAATCATTTCCACACAAATGATCTATCACCAGATGCTCGTATTTCAGGATTCCTTTGTGA 121
Db 47 CGAATCATTTCCGACAAATGATCTATTTACCAGATGCTCGTATTTCAGGATTCCTTTGTGA 106
QY 122 TAGCCGAGGGAATAATATCAATCACTTGTAGGCAATCAACAGTCCAAACGGGTATTA 181
Db 107 TAGCCGAGGGAACAATATCGATCCATTTGTTGGGCATCAACAGTCCAAACGGGTATTA 166
QY 182 ACATAGCTGGTAGAATACTAGGTGATTAGGGCTACCGTTTCTGCTGGACAATAGTAGTT 241
Db 167 ACATAGCTGGTAGAATACTAGGCGTATTGGGCGTACCGTTTCTGCTGGACAATAGTAGTT 226
QY 242 TTTATAGTTTCTTGTGTCGAATTTATGGCCCGCGGCGAGAGATCAGTGGGAAATTTTCC 301
Db 227 TTTATAGTTTCTTGTGTCGAATTTATGGCCCGCGGCGAGAGATCAGTGGGAAATTTTCC 286
QY 302 TAGAACATGTCGAAACAACTTATAATCAAAATAACAGAAATGCTTAGGAATACGGCAC 361
Db 287 TAGAACATGTCGAAACAACTTATAATCAAAATAACAGAAATGCTTAGGAATACGGCTC 346
QY 362 TTGCTCGAATCAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGATT 421
Db 347 TTGCTCGAATCAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGATT 406
QY 422 GGCTAGAAAACCGTGATGATCGAAGACGAGAGTGTCTTTATACCCCAATATATAGCTT 481
Db 407 GGCTAGAAAACCGTGATGATCGAAGACGAGAGTGTCTTTATACCCCAATATATAGCTT 466
QY 482 TAGAACTTGATTTTCTTAATCGGATGCGCTTTTCGCAATTTAGAAACCAAGAGTTCCAT 541
Db 467 TAGAACTTGATTTTCTTAATCGGATGCGCTTTTCGCAATTTAGAAACCAAGAGTTCCAT 526
QY 542 TATTAAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTAGAGAGATGCTCTC 601
Db 527 TATTAAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTAGAGAGATGCTCTC 586
QY 602 TTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATTAGCGCCCAAG 661
Db 587 TTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATTAGCGCCCAAG 646
QY 662 TGGAAACAAACGAGAGATTTATTCGCACTATTTCGCTAGAAATGCTATAATACAGTCTAAATA 721
Db 647 TGGAAACAAACGAGAGATTTATTCGCACTATTTCGCTAGAAATGCTATAATACAGTCTAAATA 706
QY 722 GCTTGAGAGGACAAAATGCGCAAGTTGGGTGCGTTATAATAATTCGTTAGAGATCTAA 781
Db 707 GCTTGAGAGGACAAAATGCGCAAGTTGGGTGCGTTATAATAATTCGTTAGAGATCTAA 766
QY 782 GCTTGAGAGGATTTAGATCTAGTGGCACTATTTCGCAAGCTATGACACTCGCACTTATCCAA 841
Db 767 GCTTGAGAGGATTTAGATCTAGTGGCACTATTTCGCAAGCTATGACACTCGCACTTATCCAA 826
QY 842 TAAATACAGTGTCTCAGTTAAACAGGGAAGTTTATACAGAGCGCAATTCGAGCAACAGGGG 901
Db 827 TAAATACAGTGTCTCAGTTAAACAGGGAAGTTTATACAGAGCGCAATTCGAGCAACAGGGG 886
QY 902 TAAATATGGCAAGTATGAATTTGGTATTAATAATAATGCACTTCTGTTTCCGCTATAGAGA 961
Db 887 TAAATATGGCAAGTATGAATTTGGTATTAATAATAATGCACTTCTGTTTCCGCTATAGAGA 946
QY 962 CTGCGGTTATCCGAAGCCCGCATCTACTTGTGATTTTCTAGAACAACTTACATTTTATAGCA 1021
Db 947 CTGCGGCTATCCGAAGCCCGCATCTACTTGTGATTTTCTAGAACAACTTACATTTTATAGCG 1006

QY 1022 CTTTCATCAGATGGAGTGTCTACTAGGCATATGACTTACTGGCGGGGCGACACAAATTCAT 1081
DB 1007 CTTTCATCAGATGGAGTAACTACTAGGCATATGACTTATTGGCGGGGCGACACGATTCAT 1066
QY 1082 CTCGGCCAAATAGGAGGGGGAATTAATACCTCAACGCATGGGTCTACCAATACTTCTATTATTA 1141
DB 1067 CTGGCCAAATAGGAGGGGGAATTAATACCTCAACGCATGGGTCTACCAATACTTCTATTATTA 1126
QY 1142 ATCCTGTAAAGATTATCAATCTCTCGAGACGCTATATTGGAGCTGAATCATATGCAAGGAG 1201
DB 1127 ATCCTGTAAATTAACGGTTCGCACTCGAGACGCTTATAGGACTGAATCATATGCAAGGAG 1186
QY 1202 TGCCTCTATGGGAATTTACCTTGAACCTTATTCATGGTCCCTACTGTTAGATTTAAAT 1261
DB 1187 TGCCTCTATGGGAATTTACCTTGAACCTTATTCATGGTCCCTACTGTTAGGTTTAAAT 1246
QY 1262 TTAGGAACCTCAGATACATTTTGAAGAGGTACTGCTAACTATAGTCAACCCCTATGAGT 1321
DB 1247 TTAGGAACCTCAGATATTTCTGTATAGGTACCGCTAACTATAGTCAACCTTATGAGT 1306
QY 1322 CACCTGGGCTTCAATTAAGATTCAAGAACTGAGAACTGAAATTAACACAGAAAACACAGAACGAC 1381
DB 1307 CACCTGGGCTTCAATTAAGATTCAAGAACTGAGAACTGAAATTAACACAGAAAACACAGAACGAC 1366
QY 1382 CAAATTAATGATCATATAGTCATAGGTATCTCACATAGGGCTCATTTCACAATCTAGGG 1441
DB 1367 CAAATTAATGATCATTTACAGTCACAGGTATCTCATATAGGTATTAATTTACAAATCCAGGG 1426
QY 1442 TGCATGTACCAGTATATTTCTTGGACGCAACCGGTAGTCAGATCGTACAAATPACCAATAGTT 1501
DB 1427 TGAATGTACCGGTATATTTCTTGGACGCACTGATGCGGACCGTACGAATPACGATTTGGAC 1486
QY 1502 CAGATAGCATAAACAAATACCAATTTGGTAAATCATTTCAACCTTAATTTCHAGGTACTCTG 1561
DB 1487 CAAATAGAAATCACCAAAATCCCAATGGTAAAGCATCCGAATCTTCTCAAGAGTACCACCTG 1546
QY 1562 TAGTCAGTGGCCAGGATTTACAGGAGGGGATATAATCCGAACCTAAAGTTAATGGTAGTG 1621
DB 1547 TTGTTAGAGGACCAAGGAATTTACTGGTGGGATATTTCTTCAAGAAACGAATPACTGGTGAT 1606
QY 1622 TACTAAGTATGGGTCTTAAATTTTAATATACATCATTTACAGCGGTATCGCGTGAGAGTTC 1681
DB 1607 TTGGACCGATTAAGAGTAACGTGTTAACGACCAATTAACACAAAGATATCGTATAGGATTC 1666
QY 1682 GTTATGCTGCTTCAAAACAAATGCTCTGAGGGTAACTGTCGAGGAGGAGTACTACTTTTG 1741
DB 1667 GCTATGCTTCAACTGTAGATTTTGAATTTCTTTGATCACGCGGAGGTACTGTAAATA 1726
QY 1742 ATCAAGGATTCCTTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCAATTTAGAT 1801
DB 1727 ATTTTAGATTTCTACGTACAAATGAAACAGTGGAGACGAACTAAATACGGAATTTTGTGA 1786
QY 1802 TTGCAGAAATTTCTGTAGGTATTAGTGCACTGCGAGTCAAACTGCTGGAATTAAGTATAA 1861
DB 1787 GACGTGCTTTTACTACACCTTTTACTTTTACACAAATTCAGAGATATAATTCGAACGTCTA 1846
QY 1862 GTAATATGCAAGGTAGACAAACG---TTTCACTTTGATTAATTAATTTGAATTCATTTCCAAATTA 1918
DB 1847 TTCAAGGCTTTAGTGGAAATGGGGAAGTGTATATAGATAAAATTTGAAATTAATTTCCAGTTA 1906
QY 1919 CTGCAACCTTCGAAGCAGAAATCAATTTAGAAAGGGGCGCAAGAGCGGTGAATCTCTGT 1978
DB 1907 CTGCACCTTCGAAGCAGAAATGATGTTAGAAAGGCGCAAGAGCGGTGAATCTCTGT 1966
QY 1979 TTACTAATACGAATCCAAAGAGAATTGAAAACAGATGTGACAGATATCATATTGATCAAG 2038
DB 1967 TTACTAATACGAATCCAAAGAGAATTGAAAACAGATGTGACAGATATCATATTGATCAAG 2026
QY 2039 TATCCAAATTTAGTGGCGTTTATCCGATGAATTTCTGCTTAGATGAAAGAGCAATTTAC 2098
DB 2027 TATCCAAATTTAGTGGATTTGTTTATCAGATGAATTTTGTCTGGATGAAAGCGAGATTTG 2086
QY 2099 TTGAGAAAGTGAAATATACGAAACGACTCAGTGTAGTAAAGAAACCTTATCTCAAGATCCAA 2158

DB 2087 CCGAGAAAGTCAACATCGGAACGACTCAGTGATGCGGGAATTTTACTTCAAGATCCAA 2146
QY 2159 ACTTCACATCCATCAATTAAGCAACCGAGACTTCATATCTACTAATGAGCAATCGAATTTCA 2218
DB 2147 ACTTCAGAGGATCAATAGAACACAGAC----- 2175
QY 2219 CATCTATCCATCAACAACTCGAACATGATGCTGGGAAGTGAGAACATTAACAATCCAGG 2278
DB 2176 -----CTGGCTGGAGGGAAGTACAGATATTACCATCCAAAG 2212
QY 2279 AAGGAAATGAGCTATTTAAAGAGAAATTACGTCACTACCGGGGACTTTTAAATGAGTGT 2338
DB 2213 GAGGAGATGAGCTATTTCAAGAGAAATTACGTCACTACCGGGTACCGTTGATGAGTGT 2272
QY 2339 ATCCGAGCTATTTTATCAAAAAATAGGAGAGTCCGAATTAAGCTTATATCTCCCTACC 2398
DB 2273 ATCCAACTGATTTTATCAGAAAAATAGATGAGTCGAAATTTAAAGCTTATATCCCGTTATG 2332
QY 2399 AATTAAGAGGTATATTGAAGATAGTCAAGATTTTAGAGATATATTGATTCGTTATAATG 2458
DB 2333 AATTAAGAGGTATATTGAAGATAGTCAAGATTTTAGAGATTTAGAAATCTATTGATCCGTTACAATG 2392
QY 2459 CGAAACATGAACAAATTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTTGAAA 2518
DB 2393 CAAACACGAAATAGTAAATGTGCCAGGCACGGGTCTCTTATGGCGGCTTTTCAGGCCAA 2452
QY 2519 GCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGGCACCAATTTTGAATCGAATCCG 2578
DB 2453 GTCCAATCGGAAGGTGCGGAGAACCGAATCGATCGGCACCACTTGAATGGAATCCG 2512
QY 2579 ATCTAGATTTGCTTCGAGAGATGAGAAATAATGTGCGCATCATTTCCCATCATTTCTCTT 2638
DB 2513 ATCTAGATTTGCTTCGAGAGACGGGGAATAATGTGCACATCATTTCCCATCATTTCACT 2572
QY 2639 TGGATATTGATATTGATGTCACAGACTTGTCATGAGAACTTAGCGGTGCGGTGATTTCA 2698
DB 2573 TGGATATTGATGTTGGATGTACAGACTTAAATGAGGACTTAGGTGTATGGTGATTTCA 2632
QY 2699 AGATTAAGACGAGGAGGTGATGCAAGACTAGGGAATCTGGAATTTATTGAGAGAGAAC 2758
DB 2633 AGATTAAGACGCAAGATGGCCATGCAAGACTAGGGAATCTAGAGTTTCTCGAAGAGAAC 2692
QY 2759 CATTTATAGAGAACGACTGTCTCGTGTGAAGAGAGCAGAGAAATAATGGAGAGACAAAC 2818
DB 2693 CATTTATAGGGAAGCACTAGCTCGTGTGAAGAGCGGAGAGAAAGTGGAGAGACAAAC 2752
QY 2819 GTGAAAACTACAATTTGGAACAAACAGAGTATATACAGAGGCAAAAGAGCTGTGATG 2878
DB 2753 GAGAGAACTGCGAGTTGGAAACAAATATTGTTTATAAAGAGGCAAAAGAAATCTGTAGATG 2812
QY 2879 CTTTATTTCTAGATTTCTCAATATATAGATTTACAAGCGGATACAACATTTGGCATGATTC 2938
DB 2813 CTTTATTTGTAACCTCTCAATATGATAGATTTACAGTGGATACGAACATCGCGATGATTC 2872
QY 2939 ATCGGCAGATAAACTTTGTTTCATCGAATTCGAGAGCTTATCTGTGAGAAATTAATCTGTTA 2998
DB 2873 ATCGGCAGATAAACCGGTTTATAGAAATCCGGAACGCTATCTGCCAGAGTTGCTGTGA 2932
QY 2999 TCCCGGTGTAAATGCGGAAATTTTGAAGAAATTAGAAGGTCCGATTAACACTGCAATCT 3058
DB 2933 TCCAGGTGTCAATGCGGCCATTTTCGAAGAAATTAGAGGACGATATTTTACAGCGTAT 2992
QY 3059 CCCTATACGATGCGAGAAATGCTGTTTAAAAATGCTGATTTTAAATAATGGAATTAGCATGCT 3118
DB 2993 CCTTATGATGCGAGAAATGCTGATTTAAAAATGCGGATTTCAATATGCTTATTATGCT 3052
QY 3119 GGAATGTAAAAGGCGATGTAGATGTACAA---CAGAGCCATACCGTTCGCTGCTGTTA 3175
DB 3053 GGAACGTGAAAGGTGATGTAGATGTAGAGAGCAAAACCAACCGTTCGCTGCTGTTA 3112
QY 3176 TCCAGAAATGGGAAGCAGAGTGTCAACAGCAGTTGCGGTCTGCTCGGGGCGGTGCTATA 3235

Db 2630 CAGACTTAATGAGGACTTAGTGTATGGTGTATATTCAGATTAAAGACGCAAGATGGCC 2689
QY 2720 ATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAACCAATTTATTAGGAGAGCACTGT 2779
Db 2690 ATGCAAGACTAGGGAATCTAGAGTTTCTGAAAGAGAAACCAATTTATTAGGGAAGCACTAG 2749
QY 2780 CTCGTGTGAAGAGAGAGAGAAATAATGGAGAGACAAACGCTGAAACCAATTTACAAATTTGAAA 2839
Db 2750 CTCGTGTGAAGAGAGAGAGAAATAATGGAGAGACAAACGCTGAAACCAATTTACAAATTTGAAA 2809
QY 2840 CAAACGAGATATATACAGAGGCAAAAGAGAGCTGTGATCTTTATTTAGATTTCTCAAT 2899
Db 2810 CAAATATTTTATAAAGAGGCAAAAGATCTGTAGATCTTTATTTAGATTTCTCAAT 2869
QY 2900 ATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCATCGGCAGAGATAAATTTGTTTC 2959
Db 2870 ATGATAGATTACAGTGGATCGAATCGCATGATTCATCGGCAGAGATAAATCGGTTTC 2929
QY 2960 ATGGAATTCGAGAGGCTTATCTGTACAGAAATTTCTGTTATCCCGGCTGTAAATTCGGAAA 3019
Db 2930 ATAGAATCCGGGAAGGCTATCTGCCAGAGTTGTCTGTGATTCAGGTGTCAATTCGGCCA 2989
QY 3020 TTTTGTGAAGAAATAGAGGTCGATATATCATCTGCAATCTCCCTATACGATGCGAGAAATG 3079
Db 2990 TTTTGTGAAGAAATAGAGGTCGATATTTTACAGCGGTATTCCTTATATGATGCGAGAAATG 3049
QY 3080 TCGTTAAAAATGTTGATTTAATAATGGAATAGCATGTGGAATGTAAAGGGCATGTAG 3139
Db 3050 TCAATTAATAATGGCAATTTCAATATAGGCTTATATGCTGGAACGTGAAGGTCAATGTAG 3109
QY 3140 ATGTACAA---CAGAGCCATCACCGTTCTGTCTGTTATCCAGAAATGGGAAGCAGAG 3196
Db 3110 ATGTAGAAGAGCAAAACACCGTTCTGCTCTGTTATCCAGAAATGGGAAGCAGAG 3169
QY 3197 TGTCAACAGCAGTTGCGTCTGTCCGGCGGTGTATATCTCCGTGTACAGCGTACA 3256
Db 3170 TGTCAACAGAGGTTGCTGTCTCCAGGTGTGTGCTGTATATCTCTGTGTACAGCATATA 3229
QY 3257 AAGAGGATATGAGAGGCTGTGTAAACGATCCATGAAATCCAGAACCAATACAGACGAAAC 3316
Db 3230 AAGAGGATATGAGAGGCTGTGTAAACGATCCATGAAATCCAGAACCAATACAGACGAAAC 3289
QY 3317 TAAATTTTAAACTGTGAAGAGAGGAGTGTATCCAAACGATACAGGAACGTGTAAATG 3376
Db 3290 TGAATTTACAGCAACTGTGTGAAGAGGAGTATATCCAAACACAGTAACTGTATA 3349
QY 3377 ATTATCTGCACACCAACAGGTACAGAGTATGT-----AATTCGGTATGCTGGATATG 3430
Db 3350 ATTATCTGGGACTCAAGAGAGAAATATGAGGGTACGTACACTTCTCGTAAATCAAGGATATG 3409
QY 3431 AGGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3490
Db 3410 AGAAGCCTATGTTAATACCTTCCGTACCGAGTATGATGATGATGATGATGATGATGATGATG 3469
QY 3491 AAACGTATACAGATGTACGAGAGATAATCATTTGTGAATATGACAGAGGATGTGGAAT 3550
Db 3470 AATCGTATACAGATGTACGAGAGATAATCATTTGTGAATATGACAGAGGATGTGGAAT 3529
QY 3551 ATCCACCACTACAGCTGTTATATGACAAAGAAATGATATCTCCAGAAACCGATA 3610
Db 3530 ACACACCACTACAGCTGTTATGTAACAAAGAAATGATATCTCCAGAGACCGATA 3589
QY 3611 AGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3670
Db 3590 AGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3649
QY 3671 TCTTTATGAGGAA 3684
Db 3650 TCTTTATGAGGAA 3663

ID XX AAS00421 standard; DNA; 3663 BP.
XX AC AAS00421;
XX DT 11-MAY-2001 (first entry)
XX DE B. thuringiensis Cry1Ba/Cry1Ca (BBC13) hybrid insecticidal toxin DNA.
XX KW Insecticidal agent; hybrid toxin; Cry1Ca; Cry1Ba; BBC13; insect control;
KW pest control; crop; insect resistant plant; transgenic plant;
XX KW crystal protein; mutant; ds.
XX OS Bacillus thuringiensis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT CDS 1.3663
FT /*tag= a
FT /partial
FT /product= "BBC13"
FT /note= "This sequence does not contain a stop codon"
XX PN WO200114562-A1.
XX PD 01-MAR-2001.
XX PF 17-AUG-2000; 2000WO-EP008042.
XX PR 19-AUG-1999; 99US-00377511.
XX PA (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX PI De Maagd RA, Bosch HJ, Carozzi NB, Warren GW;
XX DR WPI; 2001-218452/22.
XX DR P-PSDB; AAU00420.
XX PT Novel hybrid insecticidal toxin useful for controlling insects such as
PT Spodoptera exigua and Plutella xylostella, comprises domains I and II
PT from Cry1F or Cry1B toxin joined to domain III from Cry1C toxin.
XX PS Claim 6; Page 82-83; 99pp; English.
XX CC The present sequence encodes for 1 of 3 novel hybrid insecticidal toxins
CC (AAU00419-AAU00422) and is designated BBC13. BBC13 comprises domains I
CC and II from the Cry1Ba toxin at the N-terminus joined to domain III from
CC Cry1Ca toxin at the C-terminus. In the 1186 amino acid PFC1 hybrid
CC sequence, the junction between the Cry1Fa and Cry1Ca toxin domains
CC corresponds to PFC1 residues 446-454. In the 1221 amino acid BBC13 and
CC BBC15 hybrid sequences, the junction between the Cry1Ba and Cry1Ca toxin
CC domains is residues 482-488 in the BBC13 hybrid, and residues 491-494 in
CC the BBC15 hybrid. The hybrid toxins are constructed by cloning and in
CC vivo recombination. The hybrid toxins are useful for controlling an
CC insect such as Spodoptera exigua (beet armyworm), Manduca sexta (tobacco
CC hornworm), Plutella xylostella (diamondback moth), Ostrinia nubilalis
CC (European corn borer), Spodoptera frugiperda (fall armyworm), and
CC Heliothis virescens (tobacco budworm) which are major causes of
CC vegetable, fruit or ornamental flower crop losses. The nucleotide
CC sequences encoding the hybrid toxins are useful for producing an insect-
CC resistant plant by introducing the nucleotide sequence into the plant
CC preferably a transgenic plant. The nucleotide sequences can also be
CC expressed in microorganisms such as a virus, bacterium and fungus and the
CC toxin produced used as an insecticidal agent. Baculoviruses engineered
CC with the nucleotide sequence encoding a hybrid toxin can be used to
CC infect insects in vivo resulting in their death. The hybrid toxins have a
CC greater toxicity level than either of their parent toxins alone
SQ Sequence 3663 BP; 1192 A; 651 C; 828 G; 992 T; 0 U; 0 Other;

Query Match 68.9%; Score 2541.2; DB 4; Length 3663;
Best Local Similarity 82.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 548; Indels 123; Gaps 7;

QY 2 TGACTTCABAATAGAGAAAATGAGATGAAATTAATTAATGCTTTTATCGATTCCAGCTGTAT 61
Db |||||
2 TGGCTTCAAATAGAGAAAATGAGATGAAATTAATTAAT-----GCTGTAT 46
QY 62 CGAATCAATTCACACAAAATGGATCTATACAGAGTCTCGTATTTAGAGGATTCCTTTGTGTA 121
Db |||||
47 CGAATCANTTCGCACAAAATGGATCTATTAACAGATGCTCGTATTTAGGATAGCTTGTGTA 106
QY 122 TAGCCGAGGGGAATAATATCAATCCACTTTGTTAGCGATCAACAGATCCAAACGGGTATTA 181
Db |||||
107 TAGCCGAGGGGAACAATATCGATCCACTTTGTTAGCGCATCAACAGATCCAAACGGGTATTA 166
QY 182 ACATAGCTGTAGNATCTACTAGGTGTATTTAGCGCTACCGTTTGTCTGACAAATAGCTAGTT 241
Db |||||
167 ACATAGCTGTAGNATCTACTAGGTGTATTTAGCGGTATTTGGCGGTACCGTTTGTCTGACAAATAGCTAGTT 226
QY 242 TTTATAGTTTTCTTTGTTGGTGAATTTATGGCCCGCGCAGAGATCAGTGGGAAATTTTCC 301
Db |||||
227 TTTATAGTTTTCTTTGTTGGTGAATTTATGGCCCGCGCAGAGATCAGTGGGAAATTTTCC 286
QY 302 TAGAACATGTCGAACAACTTTATAAATCAACAAATAACAGAAAATGCTAGGAATACGGCAC 361
Db |||||
287 TAGAACATGTCGAACAACTTTATAAATCAACAAATAACAGAAAATGCTAGGAATACGGCTC 346
QY 362 TTGCTCGATTACAAGSTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTTGAAGATT 421
Db |||||
347 TTGCTCGATTACAAGSTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTTGAAGATT 406
QY 422 GGCTAGAAAAACCGTGATGTCGAAGAACGAGAGTGTCTTTATATACCCAATATATAGCCT 481
Db |||||
407 GGCTAGAAAACCGTGATGTCGAAGAACGAGAGTGTCTTTATATACCCAATATATAGCTT 466
QY 482 TAGAACTTGATTTCTTTAATGCGATGCGCGCTTTTCGCAATTAAGAAACCAAGAAATGCCAT 541
Db |||||
467 TAGAACTTGATTTCTTTAATGCGATGCGCGCTTTTCGCAATTAAGAAACCAAGAAATGCCAT 526
QY 542 TATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTAGAGAGATGCCTCTC 601
Db |||||
527 TATTTAGTGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTAGAGAGATGCCTCTC 586
QY 602 TTTTGTGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACCGTATTATTATGAGCGCCAAG 661
Db |||||
587 TTTTGTGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACCGTATTATTATGAGCGCCAAG 646
QY 662 TGGAAACAAACGAGAGATTTTCGACTATTCGATGCGTATGATGATATATACAGGTCTAAATA 721
Db |||||
647 TGGAAACGACGAGAGATTTTCGACTATTCGATGCGTATGATGATATATACAGGTCTAAATA 706
QY 722 GCTTGAGAGGGACAAATGCGCAAGTTGGTGGTGTATATCAATTCCTGATAGATCTAA 781
Db |||||
707 GCTTGAGAGGGACAAATGCGCAAGTTGGTGGTGTATATCAATTCCTGATAGATCTAA 766
QY 782 CGTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCAA 841
Db |||||
767 CGTTAGGAGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCAA 826
QY 842 TAAATACGATGCTCAGTTAAACGAGGAAGTTTATACAGCGCAATTTGGAGCAACAGGGG 901
Db |||||
827 TAAATACGATGCTCAGTTAAACGAGGAAGTTTATACAGCGCAATTTGGAGCAACAGGGG 886
QY 902 TAAATATGGCAAGTATGAAATTTGGTATTAATTAATGACACCTTCGTTTTCGCTATAGAGA 961
Db |||||
887 TAAATATGGCAAGTATGAAATTTGGTATTAATTAATGACACCTTCGTTCTCTGCCATAGAGG 946
QY 962 CTGCGGTATTCCGAAGCCCGCATCTACTTGATTTTCTAGAACAACTTACAAATTTTATAGCA 1021
Db |||||
947 CTGCGGTATTCCGAAGCCCGCATCTACTTGATTTTCTAGAACAACTTACAAATTTTATAGCG 1006
QY 1022 CTTTCATCAGATGAGGTGCTACTAGGCAATATGACTTACTGCGCGGGGGCACAAATTCAT 1081
Db |||||
1007 CTTTCATCAGATGAGGTGCTACTAGGCAATATGACTTACTTATGCGCGGGGGCACACGATTCAT 1066
QY 1082 CTCGGCCAAATAGGAGGGGATTAATACCTCAACGCATGGGTCTACCAATACTTCTATTA 1141

Db |||||
1067 CTCGGCCAAATAGAGGGCGGATTAATACCTCAACGCATGGGCTACCAATCTTCTATTA 1126
QY 1142 ATCTCTGTAAAGATTATCATTTCTCTCTCGAGAGCTATATTGGAGTGAATCATATGCGAGGAG 1201
Db |||||
1127 ATCTCTGTAAACATTACGGTTTCGCATCTCGAGACGTTTTATAGGACTGAATCATATGCGAGGAG 1186
QY 1202 TGCCTTCTATGGGGAATTTACCTTTGAACCTTATTATCATGGTGTCCCTACTGTTAGATTTAAT 1261
Db |||||
1187 TGCCTTCTATGGGGAATTTACCTTTGAACCTTATTATCATGGTGTCCCTACTGTTAGGTTTAAAT 1246
QY 1262 TTAGGAAACCTCAGAAATAGCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCCCTATGAGT 1321
Db |||||
1247 TTACGAAACCTCAGAAATATTTCTGTATAGAGGTACCGCTAACTATATAGTCAACCTTATGAGT 1306
QY 1322 CACCTGGGCTTCAATTTAAAGAGATTCAAGAACTGAATTTACCACAGAAAACAAACAGAACGAC 1381
Db |||||
1307 CACCTGGGCTTCAATTTAAAGAGATTCAAGAACTGAATTTACCACAGAAAACAAACAGAACGAC 1366
QY 1382 CAAAATTAATCAATATAGTCAATAGGTTATCTCAATAGGGCTCATTTTCACAATCTAGGG 1441
Db |||||
1367 CAAAATTAATCAATCTTACAGTCAAGGTTATCTCATATAGGTATAATTTTACAAATCCAGGG 1426
QY 1442 TGCATGTACAGTATATCTTGGAGCGCACCGTAGTGCAGATCGTACAAATACCAATAGTT 1501
Db |||||
1427 TGAATGTACCGGTATATCTTGGACGCATCGTAGTGCAACTCTTACAAATACAAATGATC 1486
QY 1502 CAGATAGCATAAACAAATACCAATTTGGTAAATCAATTTCAACCTTAATTTAGGTAAGT 1561
Db |||||
1487 CAGAGAGAAATTAATCAATACCTTTAGTGAAGGATTTAGAGTTTGGGGGGCGACCTCTG 1546
QY 1562 TAGTCACTGGCCAGGATTTACAGGAGGGATATTAATCCGAACATACGTTTAAATGAGT 1621
Db |||||
1547 TCATTACAGGACGAGGATTTACAGGAGGGATATCTCTCGAAGAAATACCTTTGGTGAAT 1606
QY 1622 TACTAAGTATGGTCTTAAATTTTAATAATACATCATACAGCGGTATCGCGTGAGAGTTTC 1681
Db |||||
1607 TTGTATCTCTACAGTCAATATTAATTCACCAATATACCCAAAGATACCGTTTAAAGATTTC 1666
QY 1682 GTTATGCTGCTCT-----CAAAACAATGGTCTCTGAGGGTAACTG 1720
Db |||||
1667 GTTACGCTTCCAGTAGGATGCAACGAGTTATAGTATTAACAGGAGCGCATCCACAGGAG 1726
QY 1721 TCGGAGGAGTACTACTTTTGTATCAAGGATTCCTTAGTACTATGAGTGCAATGAGTCTT 1780
Db |||||
1727 TGGAGGCGCAAGTATGTTGTAATATGCTCTTTCAGAAACTATGGAATAGGGGAGAACT 1786
QY 1781 TGACATCTCAATCATTTAGATTTGCAAGAAATTTCTGTAGGTATTAGTGATCTGGCAGTC 1840
Db |||||
1787 TAACATCTAGAACATTTAGATATACCGATTTTAGT---AATCCTTTTTCATTTAGAGCTA 1843
QY 1841 AAACCTCTGGAATAGTATTAAGTAA-----TAATGCGAGGTAGAC 1879
Db |||||
1844 ATCCAGATATAATTTGGGATAAGTGAACAACCTCTATTTGGTGCAGGTTCTTATTAGTAGCG 1903
QY 1880 AAACGTTTCACTTTGTATTAATTTGATTTCAATCCCAATTTACTGCAACTTCGAGAGCAAT 1939
Db |||||
1904 GTGAACCTTTATATAGATAAAATTTGAAATTTATTTCTAGCAGATGCAACATTTGAAGCAGAAT 1963
QY 1940 ACGATTTAGAAAGGGCGCAAGAGCGGTGAATGCTCTGTTTACTAATAAGAAATCCAAAGAA 1999
Db |||||
1964 CTGATTTAGAAAGAGCACAAAGCGGTGAATGCGCTGTTTACTTCTTCCAATCAATCG 2023
QY 2000 GATTTGAAAACAGATGTGACAGATTAATCATATTTGATCAAGTATCCAAATTTAGTGGCGTGT 2059
Db |||||
2024 GGTAAAAAACCGATGTGACGGATTAATCATATTTGATCAAGTATCCAAATTTAGTGGATGTT 2083
QY 2060 TATCGGATCAATCTCTCTTTAGATGAAAGAGAGAAATTTACTTTGAGAAAGTGAATATGCGA 2119
Db |||||
2084 TATCAGATGAATTTTCTGTTGATGAAAGAGAGATTTGTCGAGAAAGTCAAAACATGCGA 2143
QY 2120 AACGACTCAGTGATGAAAGAAACTTACTCCAAGATCCAAACTTCACATCCATCAATAAGC 2179

Db 2144 AGGACTCAGTGTAGAGCGGAATTTACTTCAAGATCCAACTTCAGAGGGATCAATAGAC 2203
Qy 2180 AACAGACTTCATATCTACTAATGAGCAATCGAATTTTCATCTATCCATGAACAATCTG 2239
Db 2204 AACCAGAC----- 2211
Qy 2240 AACATGGATGGTGGGAAGTGAACATTAATCAATCCAGGAAGGAATGACGTATTTAAAG 2299
Db 2212 --CGTGGCTGGAGAGGAAGTACAGATATTAACCATCCAGAGGAGATGACGTATTTCAAAG 2269
Qy 2300 AGAATTAAGTCTCACTACCGGGACCTTTTAATGAGTGTATCCGACGTATTTATATCAA 2359
Db 2270 AGAATTTACGTCACTACCGGGTACCGTTGATGAGTGTATCCACGTATTTATATCAGA 2329
Qy 2360 AATAGGAGAGTCGGAATTTAAAGCTTATCTCGCTACCAATTAAGAGGGTATATTGAAG 2419
Db 2330 AATAGATGAGTCGGAATTTAAAGCTTATACCGTTATGANTTAAGGGTATATCGAAG 2389
Qy 2420 ATAGTCAAGATTTAGAGATATATTTGATTCGTTAATCGGAAACATGAAACATTTGGATG 2479
Db 2390 ATAGTCAAGACTTAGAATCTATTTGATCCGTTTCAATGCAAAACACAGAAATAGTAAATG 2449
Qy 2480 TTCCAGGTACCGAGTCCGCTATGCGCGCTTTGAGTTGAAGCCCAATCGGAAGTCCGGAG 2539
Db 2450 TGCCAGGCACGGTTTCCTTATGGCGCTTTTCAGCCCAAAAGTCCAAATCGGAAGTGTGGAG 2509
Qy 2540 AACCGAATCGATCGGCACACATTTTGAATGGAATCCTGATCTAGATTTGTTCTCGCAGAG 2599
Db 2510 AACCGAATCGATCGGCACACATTTTGAATGGAATCCTGATCTAGATTTGTTCTCGCAGAG 2569
Qy 2600 ATGGAGAAATATGTCGCAATCAATCCCATCAATTTCTTTGGATATTTGATTTGGATGCA 2659
Db 2570 ACGGGGAAATATGTCACATCATTTCCCATCAATTTCCACTTTGGATATTTGATTTGGATGTA 2629
Qy 2660 CAGACTTCATCAGATCTAGCGGTGTGGTGGTATTCAGATTTAAGCGAGGAGGTGTC 2719
Db 2630 CAGACTTAATGAGGACTTAGGTGTATGGGTGATATTCAGATTTAAGCGCAAGATGGCC 2689
Qy 2720 ATGCAAGACTAGGGAATCGGAATTTATTTGAAGAGAAACCAATTTAGGAGAGCACTGT 2779
Db 2690 ATGCAAGACTAGGGAATCTAGAGTTTCTGAAAGAGAAACCAATTTAGGAGAGCACTAG 2749
Qy 2780 CTCTGTGTAAG 2839
Db 2750 CTCTGTGTAAG 2809
Qy 2840 CAAACAGAGTATATACAGAGGCAAAAGAGCTGTGGATCTTTATTTGTAGATTTCTCAAT 2899
Db 2810 CAAATATTTGTTTATAAGAGGCAAAAGAGATCTGTAGATGCTTTATTTGTAAACTCTCAAT 2869
Qy 2900 ATAAATAGATTACAGCGGATACAAACATTTGGCATGATTCATCGGCAGATAACTTTGTTTC 2959
Db 2870 ATGATAGATTACAGTGGATACGAATCCGATGATTCATCGGCAGATAAACCGGTTTC 2929
Qy 2960 ATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTGTTATCCGGGTGTAAATGCGGAAA 3019
Db 2930 ATAGAATCCGGAGAGGATCTGCCAGAGTGTCTGTGATTCAGGTGTCAATGCGGCCA 2989
Qy 3020 TTTTGAAGAAATAGAGGTGCAATTAACATCGCAATCTCCCTATACGATGCGAGAAATG 3079
Db 2990 TTTTTCGAAGAAATAGAGGACGCTATTTTACAGCGTATTTCTTTATATGATGCGAGAAATG 3049
Qy 3080 TCGTTAAATAGGTTGATTTTAAATGATTTAGCATGCTGGAATGTAAAGGCGATGTAG 3139
Db 3050 TCATTAATAAATGGCGATTTCAATAATGCGCTTATTTATGCTGGAAACGTGAAGGTCAATGTAG 3109
Qy 3140 ATGTACAA---CAGAGCCATCACCGTTCTGTCTGTTATCCAGAAATGGGAAGCAGAAG 3196
Db 3110 ATGTAGAGAGCAAAACACACCGTTTCGGTCTTGTATCCAGAAATGGGAGGCAGAAG 3169
Qy 3197 TGTCAACAGAGTTGCGCTGTGTCGGGCGGTGGCTATATCTCCGTGTACAGCGTACA 3256
Db 3170 TGTCAACAGAGTTGCGTGTCTGTCCAGGTGCTGGGTATATCTCTGTGTACAGCATATA 3229

Qy 3257 AAGAGGGATATGGAGAGGGTTGTGTAAACGATCCATGAATCGAGAAACAATACAGACGAAC 3316
Db 3230 AAGAGGGATATGGAGAGGGCTCGGTAAACGATCCATGAGATCGAAGACAATACAGACGAAC 3289
Qy 3317 TAAATTTTAAAACTGTGAAGAGGAAGTGTATCCAAACGATACAGGAACGTGTATG 3376
Db 3290 TGAATTTTCAAGCACTGTGTAGAAGAGGAAGTATATCCAAACAACAACAGTAACGTGTATA 3349
Qy 3377 ATTACTGTCACACCAAGGTACAGCATGT-----AATTTCCGTAATGCTGTATG 3430
Db 3350 ATTACTGGAAGCTCAAGAGAATATAGGGGTACGTACACTTCTCGTAATCAAGGATATG 3409
Qy 3431 AGGATGCATATCAAGTTTGATCTACAGCATCTGTGTAAATTAACAAACCGACTTATGAAGAG 3490
Db 3410 ACGAAGCCTATGGTAAATAACCTTCCGTACAGCTGATTCAGCTTCATATGAAGAA 3469
Qy 3491 AAAAGTATACAGATGTACGAAGAGATAATCAATTTGTGAATATGACAGAGGTATGTGAAT 3550
Db 3470 AATCGTATACAGATGGACGAAGAGAGAAATCTCTGTGAATCTAAACAGAGGCTATGGGAT 3529
Qy 3551 ATCCACCACTACCGCTGTTATATGACAAAGAAATTAAGAACTTCCAGAAACCGATA 3610
Db 3530 ACACACCACTACCGCTGTTATGTAAACAAAGGATTTAGAGTACTTCCAGAGACCGATA 3589
Qy 3611 AGGTATGATTTGAGATTGGAGAAACCGAAGGAAAGTTTATTTGTAGACAGCGTGAATTTAC 3670
Db 3590 AGGTATGATTTGAGATTGGAGATCCGAGAAACAGAAAGAACATTCATCGTGTAGTAGGATTTAC 3649
Qy 3671 TCCTTTATGGAGAA 3684
Db 3650 TCCTTTATGGAGAA 3663

Search completed: November 21, 2004, 09:24:38
Job time : 1588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 08:58:21 ; Search time 321 Seconds
(without alignments)
8164.108 Million cell updates/sec

Title: US-10-614-524-1

Perfect score: 3687

Sequence: 1 ttgacttcaaatagagaaaaa.....tactcttattgaggagaatag 3687

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3666.2	99.4	4173	4	US-09-661-322A-37
2	3266.6	88.6	3684	1	US-08-448-170-7
3	3266.6	88.6	3684	3	US-08-961-803-5
4	3260.2	88.4	3684	4	US-09-661-322A-62
5	3133.6	85.0	4074	1	US-08-377-690-1
6	2687.2	72.9	3934	1	US-08-100-709-3
7	2687.2	72.9	3934	1	US-08-176-865-3
8	2687.2	72.9	3934	1	US-08-474-038-3
9	2687.2	72.9	3934	2	US-08-779-046-3
10	2687.2	72.9	3934	2	US-08-881-340-3
11	1793.4	48.6	3522	1	US-08-040-751-4
12	1793.4	48.6	3522	1	US-08-291-368-1
13	1793.4	48.6	3522	2	US-08-962-190-1
14	1793.4	48.6	3522	5	PCT-US95-10310-1
15	1791.8	48.6	3522	6	Patent No. 5164180-3
16	1662	45.1	4106	1	US-08-434-823-1
17	1662	45.1	4106	1	US-08-457-366-1
18	1581.6	42.9	3558	3	US-09-178-252-22
19	1581.6	42.9	3558	4	US-09-826-660-22
20	1501	40.7	3766	1	US-08-032-364-1
21	1495.6	40.6	3624	2	US-07-951-715A-6
22	1495.6	40.6	3624	2	US-08-459-448A-6
23	1495.6	40.6	3624	3	US-08-459-595A-6
24	1495.6	40.6	3624	3	US-08-459-504B-6
25	1495.6	40.6	3624	3	US-08-459-444-6
26	1495.6	40.6	3624	3	US-09-053-549-7
27	1495.6	40.6	3624	3	US-09-547-422-6

28	1495.6	40.6	3624	4	US-09-988-462-6	Sequence 6, Appli
29	1438.6	39.0	8854	3	US-09-053-549-1	Sequence 1, Appli
30	1432.4	38.9	3713	1	US-08-100-709-1	Sequence 1, Appli
31	1432.4	38.9	3713	1	US-08-176-865-1	Sequence 1, Appli
32	1432.4	38.9	3713	1	US-08-474-038-1	Sequence 1, Appli
33	1432.4	38.9	3713	2	US-08-779-046-1	Sequence 1, Appli
34	1432.4	38.9	3713	2	US-08-881-340-1	Sequence 1, Appli
35	1411.4	38.3	4020	5	PCT-US91-02560-1	Sequence 1, Appli
36	1408.4	38.2	3522	1	US-07-828-788A-9	Sequence 9, Appli
37	1408.4	38.2	3522	1	US-08-349-867-24	Sequence 24, Appli
38	1408.4	38.2	3522	1	US-08-239-476-24	Sequence 24, Appli
39	1408.4	38.2	3522	1	US-08-356-034-7	Sequence 7, Appli
40	1408.4	38.2	3522	1	US-08-598-305A-24	Sequence 24, Appli
41	1408.4	38.2	3522	2	US-08-639-923A-24	Sequence 24, Appli
42	1408.4	38.2	3522	3	US-08-933-891-7	Sequence 7, Appli
43	1408.4	38.2	3522	3	US-09-178-252-5	Sequence 5, Appli
44	1408.4	38.2	3522	4	US-09-521-344-7	Sequence 7, Appli
45	1408.4	38.2	3522	4	US-09-826-660-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-661-322A-37

; Sequence 37, Application US/09661322A

; Patent No. 6593293

; GENERAL INFORMATION:

; APPLICANT: Baum, James A.

; APPLICANT: Chu, Chih-Rei

; APPLICANT: Donovan, William P.

; APPLICANT: Gilmer, Amy J.

; APPLICANT: Rupar, Mark J.

; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Compos

; FILE REFERENCE: MECO201

; CURRENT APPLICATION NUMBER: US/09/661.322A

; CURRENT FILING DATE: 2000-09-13

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: Patent version 3.0

; SEQ ID NO 37

; LENGTH: 4173

; TYPE: DNA

; ORGANISM: Bacillus thuringiensis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(3687)

US-09-661-322A-37

Query Match	99.4%	Score 3666.2;	DB 4;	Length 4173;
Best Local Similarity	99.6%	Pred. No. 0;		
Matches 3674;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;
Qy	1	TTGACTTCAAATAGGAAAATGAGAATGAAATATAAATGCTTTATCGATTCCAGCTGTA	60	
Db	1	TTGACTTCAAATAGGAAAATGAGAATGAAATATAAATGCTTTATCGATTCCAGCTGTA	60	
Qy	61	TCGAATCATTCACACAAATCGATCTATCACCAGATGCTCGTATTGAGGATTCCTTTGTGT	120	
Db	61	TCGAATCATTCACACAAATCGATCTATCACCAGATGCTCGTATTGAGGATTCCTTTGTGT	120	
Qy	121	ATAGCCGAGGGGAATAATATCAATCCCTTCTTAGCGCATCAACAGTCCAAACGGGTATT	180	
Db	121	ATAGCCGAGGGGAATAATATCAATCCCTTCTTAGCGCATCAACAGTCCAAACGGGTATT	180	
Qy	181	AACATAGCTGTAGATACTAGGTGTATTAGGCGTACCGTTTGTGGAACAATAGCTAGT	240	
Db	181	AACATAGCTGTAGATACTAGGTGTATTAGGCGTACCGTTTGTGGAACAATAGCTAGT	240	
Qy	241	TTTTATAGTTTCTTTGTTGGTAATTATGGCCCGCGGCAGAGATCATGTGGGAATTTTC	300	
Db	241	TTTTATAGTTTCTTTGTTGGTAATTATGGCCCGCGGCAGAGATCATGTGGGAATTTTC	300	

Qy	301	CTAGAACATGCTCGAACAACTTATAAATCAACAATAACAGAAATGCTAGGAATACGGCA	360
Db	301	CTAGAACATGCTCGAACAACTTATAAATCAACAATAACAGAAATGCTAGGAATACGGCA	360
Qy	361	CTTGCTCGAATACAAAGGTTAGAGATTCCTTTAGAGCCCTATCAACAGTCACCTTGAAGAT	420
Db	361	CTTGCTCGAATACAAAGGTTAGAGATTCCTTTAGAGCCCTATCAACAGTCACCTTGAAGAT	420
Qy	421	TGCGTAGAAAACCGTGATGATCAAGAACGAGAAAGTGTTCTTTATACCCCAATATATAGCC	480
Db	421	TGCGTAGAAAACCGTGATGATCAAGAACGAGAAAGTGTTCTTTATACCCCAATATATAGCC	480
Qy	481	TTAGAACTCATTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAAAGTTCCA	540
Db	481	TTAGAACTCATTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAAAGTTCCA	540
Qy	541	TTATTAAATGATATGCTCAAGCTGCAAAATTTACACTAATTAATTTAGAGAGATGCTCT	600
Db	541	TTATTAAATGATATGCTCAAGCTGCAAAATTTACACTAATTAATTTAGAGAGATGCTCT	600
Qy	601	CTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACAGTTATTATGAGCGCCAA	660
Db	601	CTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACAGTTATTATGAGCGCCAA	660
Qy	661	GTGGAACAAACGAGAGATTAATCCGACTATTGGGTAGAAATGGTATTAATACAGGTCTAAAT	720
Db	661	GTGGAACAAACGAGAGATTAATCCGACTATTGGGTAGAAATGGTATTAATACAGGTCTAAAT	720
Qy	721	AGCTTGAGAGGGCAAAATGCGCAAGTTGGGTGGTTAATCAATTCGGTAGAGATCTA	780
Db	721	AGCTTGAGAGGGCAAAATGCGCAAGTTGGGTGGTTAATCAATTCGGTAGAGATCTA	780
Qy	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTTACGACTCGCACTTATCCA	840
Db	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTTACGACTCGCACTTATCCA	840
Qy	841	ATAAATACGAGTGCTCAGTTTAAACAAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG	900
Db	841	ATAAATACGAGTGCTCAGTTTAAACAAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG	900
Qy	901	GTAATAATGCGAAGTATGAATTTGGTATAATAATAATGCACTTCGTTTCCGCTATAGAG	960
Db	901	GTAATAATGCGAAGTATGAATTTGGTATAATAATAATGCACTTCGTTTCCGCTATAGAG	960
Qy	961	ACTCGGTTATCCGAAGCCGCACTACTTGATTTTCTAGAACAACTTACAAATTTTATAGC	1020
Db	961	ACTCGGTTATCCGAAGCCGCACTACTTGATTTTCTAGAACAACTTACAAATTTTATAGC	1020
Qy	1021	ACTTCATCAGATGGAGTGCTACTAGGCATATGACTTACTGGGGGGGCGCACAAATTCAA	1080
Db	1021	ACTTCATCAGATGGAGTGCTACTAGGCATATGACTTACTGGGGGGGCGCACAAATTCAA	1080
Qy	1081	TCTCGGCCAATPAGAGGCGGAATTAATACTCAACGCAATGGGTCTACAAATCTTCTATT	1140
Db	1081	TCTCGGCCAATPAGAGGCGGAATTAATACTCAACGCAATGGGTCTACAAATCTTCTATT	1140
Qy	1141	AATCCTGTAGATTAATCATTTCTCTCGAGACGTATATGGACTGAACTCATATGCGAGGA	1200
Db	1141	AATCCTGTAGATTAATCATTTCTCTCGAGACGTATATGGACTGAACTCATATGCGAGGA	1200
Qy	1201	GTGCTTCTATGGGAAATTTACTCTTGAACCTATTCAATGGTGTCCCTACTGTTAGATTTAAT	1260
Db	1201	GTGCTTCTATGGGAAATTTACTCTTGAACCTATTCAATGGTGTCCCTACTGTTAGATTTAAT	1260
Qy	1261	TTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCCCTATGAG	1320
Db	1261	TTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCCCTATGAG	1320
Qy	1321	TCACCTGGCTTCAATTTAAAGATTCAGAACTGAAATTTACCACAGAAACCAACAGAACGA	1380
Db	1321	TCACCTGGCTTCAATTTAAAGATTCAGAACTGAAATTTACCACAGAAACCAACAGAACGA	1380
Qy	1381	CCAAATTATGAATCATATAGTCATAGGTTATCTCACATAGGGCTCATTTTCAAAATCTAGG	1440

Db	1381	CCAAATTATGAATCATATAGTCATAGGTTATCTCACATAGGGCTCATTTTCAAAATCTAGG	1440
Qy	1441	GTGCAATGACCAAGTATATTTCTTGGACGCACCGTAGTCAGATCGTACAAATACCAATTAGT	1500
Db	1441	GTGCAATGACCAAGTATATTTCTTGGACGCACCGTAGTCAGATCGTACAAATACCAATTAGT	1500
Qy	1501	TCAGATAGCATAAACAAATACCATTTGGTAAATCATTTCAACCTTTAAATTCAGGTACCTCT	1560
Db	1501	TCAGATAGCATAAACAAATACCATTTGGTAAATCATTTCAACCTTTAAATTCAGGTACCTCT	1560
Qy	1561	GTAGTCAGTGGCCCAAGGATTTACAGAGGGGATATAATCCCGAACTAAACGTTAATGGTAGT	1620
Db	1561	GTAGTCAGTGGCCCAAGGATTTACAGAGGGGATATAATCCCGAACTAAACGTTAATGGTAGT	1620
Qy	1621	GTACTAAATGATGGGTCTTAAATTTAATATCATCATTTACAGCGGTATCGCGTAGAGATT	1680
Db	1621	GTACTAAATGATGGGTCTTAAATTTAATATCATCATTTACAGCGGTATCGCGTAGAGATT	1680
Qy	1681	CGTTATGCTGCTCTCAAAATGCTCTGAGGGTAACTGTCGAGGGAGTAGTACTACTTTT	1740
Db	1681	CGTTATGCTGCTCTCAAAATGCTCTGAGGGTAACTGTCGAGGGAGTAGTACTACTTTT	1740
Qy	1741	GATCAAGGATTCCTTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCATTTTGA	1800
Db	1741	GATCAAGGATTCCTTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCATTTTGA	1800
Qy	1801	TTTGCAGATTTTCTGTAGGTATTAGTCATCTGCGAGTCAAACTGCTCGAATAAGTATA	1860
Db	1801	TTTGCAGATTTTCTGTAGGTATTAGTCATCTGCGAGTCAAACTGCTCGAATAAGTATA	1860
Qy	1861	AGTAATAATGCAAGTAGACAAACGTTTCACTTTGATATAAATTTGAAATTCATTTCAATTACT	1920
Db	1861	AGTAATAATGCAAGTAGACAAACGTTTCACTTTGATATAAATTTGAAATTCATTTCAATTACT	1920
Qy	1921	GCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGAGCGGTGATGCTCTGTTT	1980
Db	1921	GCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGAGCGGTGATGCTCTGTTT	1980
Qy	1981	ACTAATACCAATCCCAAGAGATTGAAACAGAGATGACAGATTATCATATTGATCAAGTA	2040
Db	1981	ACTAATACCAATCCCAAGAGATTGAAACAGAGATGACAGATTATCATATTGATCAAGTA	2040
Qy	2041	TCCAAATTTAGTGGCGTGTATTCGAGTGAATTTCTGTTAGATGAAAGAGAGAAATTAATT	2100
Db	2041	TCCAAATTTAGTGGCGTGTATTCGAGTGAATTTCTGTTAGATGAAAGAGAGAAATTAATT	2100
Qy	2101	GAGAAAGTGAATATGCGGAAAACGACTCAGTGATGAAAGAAACTTACTCCAAGATCCAAAC	2160
Db	2101	GAGAAAGTGAATATGCGGAAAACGACTCAGTGATGAAAGAAACTTACTCCAAGATCCAAAC	2160
Qy	2161	TTCCATCCATCAATAAGCAACCCAGACTTTCATATCTTAATGAGCAATTCGAAATTTTCA	2220
Db	2161	TTCCATCCATCAATAAGCAACCCAGACTTTCATATCTTAATGAGCAATTCGAAATTTTCA	2220
Qy	2221	TCTATCCATGAAACAACTCTGAACATGATGGTGGGGAGTGAAGAACATTAACAATCCAGAA	2280
Db	2221	TCTATCCATGAAACAACTCTGAACATGATGGTGGGGAGTGAAGAACATTAACAATCCAGAA	2280
Qy	2281	GGAATGACGTATTTAAAGAGAAATACGTCACACTACCGGGGACTTTTAAATGAGTGTAT	2340
Db	2281	GGAATGACGTATTTAAAGAGAAATACGTCACACTACCGGGGACTTTTAAATGAGTGTAT	2340
Qy	2341	CCGACGTTATTTATCAAAAATAGGAGAGTCCGAATTTAAAGCTTTATCTCCGCTACCAA	2400
Db	2341	CCGACGTTATTTATCAAAAATAGGAGAGTCCGAATTTAAAGCTTTATCTCCGCTACCAA	2400
Qy	2401	TTAAGAGGGTATTTGAAGATAGTCAAGATTTTGAAGATATTTTGAATTCGTTAATGCG	2460
Db	2401	TTAAGAGGGTATTTGAAGATAGTCAAGATTTTGAAGATATTTTGAATTCGTTAATGCG	2460
Qy	2461	AAACATGAACATTTGATGTTCCAGGTACCGGTCCGTATGGCGCTTTCCAGTTGAAGC	2520

Db 2461 AAACATGAAACATTGGATGTTCCAGGTACCGAGTCCGATATGGCCGCTTTCAGTTGAAAGC 2520
Qy 2521 CCAATCGAAGGTGCGGAGAACCGAATCGATGCGCACCAACATTTTGAATGGAATCCTGAT 2580
Db 2521 CCAATCGAAGGTGCGGAGAACCGAATCGATGCGCACCAACATTTTGAATGGAATCCTGAT 2580
Qy 2581 CTAGATTGTTCTCGCAGAGATGGAGAAATATGTGCGCATCATTCGCCATCATTTCTCTTTG 2640
Db 2581 CTAGATTGTTCTCGCAGAGATGGAGAAATATGTGCGCATCATTCGCCATCATTTCTCTTTG 2640
Qy 2641 GATATTGATATTGGATGCACAGACTTGCATGAGATCTAGGCGTGGTGGTATTCAAG 2700
Db 2641 GATATTGATATTGGATGCACAGACTTGCATGAGATCTAGGCGTGGTGGTATTCAAG 2700
Qy 2701 ATTAGAGCGCAGGAAGGTGATGCAAGACTTAGGGAATCTGGAATTTATTTGAAGAGAAACCA 2760
Db 2701 ATTAGAGCGCAGGAAGGTGATGCAAGACTTAGGGAATCTGGAATTTATTTGAAGAGAAACCA 2760
Qy 2761 TTATTAGAGAGCACTGTCTGCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Db 2761 TTATTAGAGAGCACTGTCTGCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Qy 2821 GAAAACTACAAATTTGGAACAAACGAGTATATACAGAGGCAAAAGAGAGCTGTGGATGCT 2880
Db 2821 GAAAACTACAAATTTGGAACAAACGAGTATATACAGAGGCAAAAGAGAGCTGTGGATGCT 2880
Qy 2881 TTATTGTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
Db 2881 TTATTGTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
Qy 2941 GCGGAGAGATAAATTGTTTCATCGAATTCGAGAGGCTTATCTGTCAAGATTTATCTGTTATC 3000
Db 2941 GCGGAGAGATAAATTGTTTCATCGAATTCGAGAGGCTTATCTGTCAAGATTTATCTGTTATC 3000
Qy 3001 CCGGCTGTAATCGGGAATTTTCAAGATTTAGAGGTCGCATTTACCTGCAATCTCC 3060
Db 3001 CCGGCTGTAATCGGGAATTTTCAAGATTTAGAGGTCGCATTTACCTGCAATCTCC 3060
Qy 3061 CTATACGATGCGAGAAATGTCCTTAAAAATGCTGATTTTAAATAGGATTAGCATGCTGG 3120
Db 3061 CTATACGATGCGAGAAATGTCCTTAAAAATGCTGATTTTAAATAGGATTAGCATGCTGG 3120
Qy 3121 AATGTAAGGCGCATGTAGATGTACACAGAGCCATCACCGTTCTGCTCTTGTATTCCCA 3180
Db 3121 AATGTAAGGCGCATGTAGATGTACACAGAGCCATCACCGTTCTGCTCTTGTATTCCCA 3180
Qy 3181 GAATGGGAGCAGAGTGTACAGAGCTTGCCTGCTGTCGGGCGGCTATATCTC 3240
Db 3181 GAATGGGAGCAGAGTGTACAGAGCTTGCCTGCTGTCGGGCGGCTATATCTC 3240
Qy 3241 CGTGTACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAATCGAG 3300
Db 3241 CGTGTACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAATCGAG 3300
Qy 3301 AACATACAGACGAATCTAAAAATTTAAAACTGTGGAAGAGGAGTGTATCCACGGAT 3360
Db 3301 AACATACAGACGAATCTAAAAATTTAAAACTGTGGAAGAGGAGTGTATCCACGGAT 3360
Qy 3361 ACAGGAGCGTGAATGATTATCTCCACCAAGGTACAGCAGTGTGAATCCCGTAAT 3420
Db 3361 ACAGGAGCGTGAATGATTATCTCCACCAAGGTACAGCAGTGTGAATCCCGTAAT 3420
Qy 3421 GCTGATATGAGATGCAATGAACTGATCTACAGCATCTGTTAAATTAACAACCGACT 3480
Db 3421 GCTGATATGAGATGCAATGAACTGATCTACAGCATCTGTTAAATTAACAACCGACT 3480
Qy 3481 TATGAAGAGAAACGTATACAGATGTACGAAGAGATAATCAATTTGGAATATGACAGAGGG 3540
Db 3481 TATGAAGAGAAACGTATACAGATGTACGAAGAGATAATCAATTTGGAATATGACAGAGGG 3540
Qy 3541 TATGTGAATTTCCACCACTACCAGCTGTTATATGACAAAAGAAATTAAGATTAATCTCCA 3600
Db 3541 TATGTGAATTTCCACCACTACCAGCTGTTATATGTAAGAAAAGAAATTAAGATTAATCTCCA 3600

Qy 3601 GAAACCGATAGGTATGATTGAGATTGGAGAAACCGAAGGAAAGTTTATTGTAGACAGC 3660
Db 3601 GAAACCGATAGGTATGATTGAGATTGGAGAAACCGAAGGAAAGTTTATTGTAGACAGC 3660
Qy 3661 GTGGAATTACTCTCTTATGAGAGAAATAG 3687
Db 3661 ATAGAATTACTCTCTTATGGAAGAAATAG 3687

RESULT 2

US-08-448-170-7
; Sequence 7, Application US/08448170
; Patent No. 5723758
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Steiman, Steve
; TITLE OF INVENTION: No. 5723758el Bacillus thuringiensis Isolate Denoted
; TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,170
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: W/S 102D.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3684 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-448-170-7

Query Match 88.6%; Score 3266.6; DB 1; Length 3684;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 3449; Conservative 0; Mismatches 229; Indels 15; Gaps 3;

Qy 1 TTGACTTCAATAGAAAAATGAGAAATGAGAAATTAATAATGCTTTTATCGATTCCAGCTGTA 60
Db 1 TTGACTTCAATAGAAAAATGAGAAATTAATAATGCTTTTATCGATTCCAGCTGTA 60
Qy 61 TCGAATCATTTCCACAAATGGATCTATCACCAGATGCTCGTATTTGAGGATTTCTTTGTGT 120

2275 CAGGAAGGAATGACGTATTTAAAGAGAAATTACGTACACTACCGGAGACTTTTAAATGAG 2334
2272 CAGGAAGGAATGACGTATTTAAAGAGAAATTACGTACACTACCGGAGACTTTTAAATGAG 2331
2335 TGTATCCGACGTATTTATATCAAAAAATAGAGAGTCGGAAATTAAGCTTATATCTGCG 2394
2332 TGTATCCGACGTATTTATATCAAAAAATAGGGGAGCGGAAATTAAGCTTATATCTGCG 2391
2395 TACCAATTAAGAGGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2454
2392 TACCAATTAAGTGGCTATATTGAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2451
2455 AATCGAAACATGAACATTGATGTTCCAGTACCGAGTCGGATGCGCGCTTTTCAGTT 2514
2452 AATCGAAACATGAACATTGATGTTCCAGTACCGAGTCGGATGCGCGCTTTTCAGTT 2511
2515 GAAAGCCCAATCGGAGGTGCGGAGAACCGAATCGATCGACCACTTTTGAATGGAAT 2574
2512 GAAAGCCCAATCGGAGGTGCGGAGAACCGAATCGATCGACCACTTTTGAATGGAAT 2571
2575 CCTGATCTAGATGTTCTTCGACAGATGAGAGAAATGTGCGCATCATTTCCCATCATTTTC 2634
2572 CCTGATCTAGATGTTCTTCGACAGATGAGAGAAATGTGCGCATCATTTCCCATCATTTTC 2631
2635 TCTTTGGATATTGATTTGGATGCAAGACTTTGCAATGAGAAATCTAGGCGTGTGGTGGTA 2694
2632 TCTTTGGATATTGATTTGGATGCAAGACTTTGCAATGAGAAATCTAGGCGTGTGGTGGTA 2691
2695 TTCAAGATTAAAGCAGGAGGTCATGCAAGACTAGGAAATCTGGAATTTATTTGAAGAG 2754
2692 TTCAAGATTAAAGCAGGAGGTCATGCAAGACTAGGAAATCTGGAATTTATTTGAAGAG 2751
2755 AAACCAATTATTAGGAGACACTGCTCGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAC 2814
2752 AAACCAATTATTAGGAGACACTGCTCGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAC 2811
2815 AAACGTGAAAACTCAATTGGAACAAAAACAGAGTATATACAGAGGCAAGAAAGCTGTG 2874
2812 AAACGTGAAAACTCAATTGGAACAAAAACAGAGTATATACAGAGGCAAGAAAGCTGTG 2871
2875 GATGCTTTATTTGTAGATTCTCAATATATAGATTACAGCGGATACAAACATTTGGCATG 2934
2872 GATGCTTTATTTGTAGATTCTCAATATATAGATTACAGCGGATACAAACATTTGGCATG 2931
2935 ATTCAATGCGGAGATAACTGCTTCATCGAATTCGAGAGGCTTATCTGTGCAAGATTATCT 2994
2932 ATTCAATGCGGAGATAACTGCTTCATCGAATTCGAGAGGCTTATCTGTGCAAGATTATCT 2991
2995 GTTATCCCGGTTAAATCGGAAATTTTGAAGAAATTAGAGGTGCGCAATTATCACTGCA 3054
2992 GTTATCCCGGTTAAATCGGAAATTTTGAAGAAATTAGAGGTGCGCAATTATCACTGCA 3051
3055 ATCTCCCTATACGATCGGAGAAATGTCGTTAAATAATGTTGATTTTAAATAGGATAGCA 3114
3052 ATCTCCCTATACGATCGGAGAAATGTCGTTAAATAATGTTGATTTTAAATAGGATAGCA 3111
3115 TGCTGGAATGTAAGAGGTCATGATGTAACAGAGCCATCACCGTTCTGTCCTCTGTT 3174
3112 TGCTGGAATGTAAGAGGTCATGATGTAACAGAGCCATCACCGTTCTGTCCTCTGTT 3171
3175 ATCCCAAGATGGAAGAGCAGAAAGTGTCAAGAGGATTCGCTGTGTCGGGGGCGTGGCTAT 3234
3172 ATCCCAAGATGGAAGAGCAGAAAGTGTCAAGAGGATTCGCTGTGTCGGGGGCGTGGCTAT 3231
3235 ATCTCCCTGTCACAGGTCACAAAGGATATGAGAGGGTGTGTAAACGATCCATGAA 3294
3232 ATCTCCCTGTCACAGGTCACAAAGGATATGAGAGGGTGTGTAAACGATCCATGAA 3291
3295 ATCGAGAACATACAGAGCACTAAATTTTAAACTGTGAGAGAGAGAGAGTGTATCCA 3354
3292 ATCGAGAACATACAGAGCACTAAATTTTAAACTGTGAGAGAGAGAGAGTGTATCCA 3351

3355 ACGGATACAGGAACGTGTAATGATTATATCTGACACCAAGGTACAGAGTATGTATTC 3414
3352 ACGGATACAGGAACGTGTAATGATTATATCTGACACCAAGGTACAGAGTATGTATTC 3411
3415 CGTAATGCTGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTTAATTACAAA 3474
3412 CGTAATGCTGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTTAATTACAAA 3471
3475 CCGACTTATGAAGAGAAACGTATACAGATGTACGAAGAGATAATCATTTGTGAATATGAC 3534
3472 CCGACTTATGAAGAGAAACGTATACAGATGTACGAAGAGATAATCATTTGTGAATATGAC 3531
3535 AGAGGATATGAAATTTATCCACCACTACCACTGGTTATATGACAAAAAGAAATTAGAATAC 3594
3532 AGAGGATATGAAATTTATCCACCACTACCACTGGTTATATGACAAAAAGAAATTAGAATAC 3591
3595 TTCCAGAGAAACCGATAGGATGATGAGATTGAGAAACCGAGAGGAGTTTATTGTA 3654
3592 TTCCAGAGAAACCGATAGGATGATGAGATTGAGAAACCGAGAGGAGTTTATTGTA 3651
3655 GACAGCGTGGAAATTAATCTCTTATGAGGAATAG 3687
3652 GACAGCGTGGAAATTAATCTCTTATGAGGAATAG 3684

RESULT 3

US-08-961-803-5
; Sequence 5, Application US/08961803
; Patent No. 6150589
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stelman, Steve
; TITLE OF INVENTION: No. 6150589el Bacillus thuringiensis Isolate Denoted
; TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jay M. Sanders
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,803
; FILING DATE: 31-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/448,170
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: M/S 102DCD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100

```
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3684 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-961-803--5

Query Match      88.6%; Score 3265.6; DB 3; Length 3684;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 3449; Conservative 0; Mismatches 229; Indels 15; Gaps 3;

QY      1 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATATATAAATGCTTTATCGAATCCAGCTGTA 60
DB      1 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATATATAAATGCTTTATCGAATCCAGCTGTA 60

QY      61 TCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTGAGGATTCCTTTGTGT 120
DB      61 TCGAATCATTTCCGACAAATGAATCTATCAACGATGCTCGTATTGAGGATGCTTTGTGT 120

QY      121 ATAGCCGAGGGGAATATATCAATCCACTTGTAGGCGCATCAACAGTCCAAACGGGTATT 180
DB      121 ATAGCCGAGGGGAATATATCGATCCATTTGTTAGGCGCATCAACAGTCCAAACGGGTATT 180

QY      181 AACATAGCTGGTGAATACATAGGTGATTTAGGCGTACCGTTTCTCGGACAAATPAGTAGT 240
DB      181 AACATAGCTGGTGAATACATAGGTGATTTAGGCGTACCGTTTCTCGGACAAATPAGTAGT 240

QY      241 TTTTATAGTTTCTCTGTTGGTGAATATATGCGCGCGGCGGAGAGATCAGTGGGAAATTTTC 300
DB      241 TTTTATAGTTTCTCTGTTGGTGAATATATGCGCGCGGCGGAGAGATCCTTGGGAAATTTTC 300

QY      301 CTAGAACATGTCGAAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
DB      301 CTAGAACATGTCGACAACTTATAAGNACAAAGTAACAGAAATATCTAGGATACGGCT 360

QY      361 CTTGCTCGATTACAAGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTTGAAGAT 420
DB      361 CTTGCTCGATTACAAGTTTAGGAAATTCCTTTAGAGCCTATCAACAGTCACTTTGAAGAT 420

QY      421 TGCTAGAAAACCGTGATGATCAAGAACGAGAGAGTCTTTATACCCCAATATATAGCC 480
DB      421 TGCTAGAAAACCGTGATGATCAAGAACGAGAGAGTCTTTATACCCCAATATATAGCC 480

QY      481 TTAGAACTTGATTTCTTAATGCGATGCGCCTTTTCGCAATTAGAAAACCAAGAAATTCCA 540
DB      481 TTAGAACTTGATTTCTTAATGCGATGCGCCTTTTCGCAATTAGAAAACCAAGAAATTCCA 540

QY      541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACTATATATTTAGAGAGATGCTCTCT 600
DB      541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACTATATATTTAGAGAGATGCTCTCT 600

QY      601 CTTTGTGTTAGTGAATTTGGGCTTACATCCGAGGAAATCAACGTTATTATGAGCGCCAA 660
DB      601 CTTTGTGTTAGTGAATTTGGGCTTACATCCGAGGAAATCAACGTTATTATGAGCGCCAA 660

QY      661 GTGAAACAAACGAGAGATTTATCCGACTATTGGGTAGAAATGGGTATATATACAGTCTAAAT 720
DB      661 GTGAAACAAACGAGAGATTTATCTGATTTATGGCAAGATGGTATATACGGGTTAAAT 720

QY      721 AGCTTGAGGGGACAAATGCGCAAGTTGGGTGGTGTATTAATCAATTCGTTAGAGATCTA 780
DB      721 AATTTGAGGGGACAAATGCTGAAAGTTGGTTGCGATATAATCAATTCGTTAGAGACTTA 780

QY      781 ACGTTAGGGGTATATAGTCTAGTGGCACTATTCGAGCTATGACACTCCACTTATCCA 840
DB      781 ACGTTAGGGGTATATAGTCTAGTGGCACTATTCGAGCTATGACAGCGGTGTTATCCA 840

QY      841 ATAAATACGAGTCTCAGTTTAAACAAAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG 900
DB      841 ATGAATACCAAGTCTCAATTTAAACAGAGAAATTTATACAGATCCAAATTTGGGAGAACAAAT 900

901 GTAAAT-----ATGCAAGTATGAATGGTATATAATAATGCACTTCGTTTCCGCT 954
901 GCACCTTCAGGATTTGCAAGTACGAATGGTTAATAATAATGCACTTCGTTTCCGCT 960

955 ATAGAGACTGCGGTTATCCGAAGCCCGCATCTACTTTGATTTTCTAGAAACAACTTACAAT 1014
961 ATAGAGGCTGCGGTTATTTAGGCTCGGATCTACTTTGATTTTCCAGAACAGCTTACAAT 1020

1015 TTTAGCACTTCATCAGATGGAGTGTACTAGGCATATAGCTTACTTGGCGGGGACACACA 1074
1021 TTTAGCGTATTAACTGATGGAGTAACTCAATATATGAATTTACTTGGGTGGGACATAGA 1080

1075 ATTCATCTCGCGCCCAATAGGAGCGGATTAATCTCAACGCATGCTCTACCAATATCT 1134
1081 CTTGAATCGGAAACAAATAGGGGGCTTTAAGTACTCTCGACACACGGAATACCAATATCT 1140

1135 TCTATTAAATCTGTGAAGATTATCTTCTCTCGAGAGCTATATTTGGACTGAATCATAT 1194
1141 TCTATTAAATCTGTGAAGATTATCTTCTCTCGAGAGCTTATAGACAGATCATTT 1200

1195 GCAGAGTGTCTTATGCGGAAATTTACCTTGAACTTATCTATGCTGCTCTCTGTTAGA 1254
1201 GCAGGATATAAT-----ACTTCTAACTACTCTCTGTGAATGGAGTACTCTTGGGCTAGA 1254

1255 TTTAATTTTAGGAAACCTCGAGAAATTTTGAAGAGGTACTGCTTAATATATAGTCAACCC 1314
1255 TTTAATTTGAGAAATCCCTGAAATTTCTCT---AGAGGTAGCTCTCTCTATATATAGG 1311

1315 TATGAGTCACTGGGCTTCAATTTAAAGATTCAGAAACTGAAATACCAACAGAAACAAACA 1374
1312 TATCTGGAGTGGGACACACTATTTGATTCAGAAACTGAAATACCAACAGAAACAAACA 1371

1375 GAAACGCAAAATTTAATGATCATATATAGTATAGTGTATCTCAATAGGGGTCTCAATTTCA 1434
1372 GAAACGCAAAATTTAATGATCTTACAGTATAGTATCTAATATAAGACTAATATATCAGGA 1431

1435 TCTAGGGTGCATGTACCAGTATATTTTGGAGCGACCGTAGTGCAGATCGTACAAATACC 1494
1432 AACACTTTGAGAGACCGAGTATATTTTGGAGCGACCGTAGTGCAGATCGTACAAATACC 1491

1495 ATTAGTTTCAGATAGATTAACAAATATCCATTTGGTAAAAATCAATTCACCTTAATTCAGGT 1554
1492 ATTAGTTTCAGATAGATTAACAAATATCCATTTGGTAAAAATCAATTCACCTTAATTCAGGT 1551

1555 ACCTCTGATGTCAGTGGCCAGGATTTACAGAGGGGATATAATCCGAATCTAACCGTTAAT 1614
1552 ACCTCTGATGTCAGTGGCCAGGATTTACAGAGGGGATATAATCCGAATCTAACCGTTAAT 1611

1615 GGTAGTGTACTAAGTATGGTCTTAAATTTTAAATATACATCATTTACAGCGGTATCCGGTG 1674
1612 GGTAGTGTACTAAGTATGGTCTTAAATTTTAAATATACATCATTTACAGCGGTATCCGGTG 1671

1675 AGAGTTTCGTTATGCTGCTTCTCAAAACAAATGGTCTCTGAGGGTAACTGTCGAGGGGAGTACT 1734
1672 AGAGTTTCGTTATGCTGCTTCTCAAAACAAATGGTCTCTGAGGGTAACTGTCGAGGGGAGTACT 1731

1735 ACTTTTGAATCAAGGATTCCTAGTACTAGTGCAGAAATAGTCTTTTGACATCTCAATCA 1794
1732 ACTTTTGAATCAAGGATTCCTAGTACTAGTGCAGAAATAGTCTTTTGACATCTCAATCA 1791

1795 TTTAGATTTGCGAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA 1854
1792 TTTAGATTTGCGAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA 1851

1855 AGTATAAGTAAATAATGCGAGTACAAACGTTTCACTTTTGATAAAAAATTTGAATTCATTTCCA 1914
1852 AGTATAAGTAAATAATGCGAGTACAAACGTTTCACTTTTGATAAAAAATTTGAATTCATTTCCA 1911

1915 ATTTACTGCAACCTTTCGAAAGCAAAATACGATTTAGAAAGCGGCAAGCGCGGTGAATGCT 1974
1912 ATTTACTGCAACCTTTCGAAAGCAAAATATGATTTAGAAAGCGGCAAGCGCGGTGAATGCT 1971
```

QY 1975 CTGTTTACTAATACGAATCCAAAGAGATTGAAAACAGATGTGACAGATTATCATATTGAT 2034
DB 1972 CTGTTTACTAATACGAATCCAAAGAGATTGAAAACAGATGTGACAGATTATCATATTGAT 2031
QY 2035 CAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATTCGTCTAGATGAAAAGAGAGAA 2094
DB 2032 GAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATTCGTCTAGATGAAAAGAGAGAA 2091
QY 2095 TTACTTTGAGAAAGTGAATATATCGGAAACGACTCAGTGTGAAAGAACTTACTCCAAGAT 2154
DB 2092 TTACTTTGAGAAAGTGAATATATCGGAAACGACTCAGTGTGAAAGAACTTACTCCAAGAT 2151
QY 2155 CAAACTTTCACATCCATCAATTAAGCAACAGACTTTCATATCTACTTAATGAGCAATFCGAAT 2214
DB 2152 CAAACTTTCACATCCATCAATTAAGCAACAGACTTTCATATCTACTTAATGAGCAATFCGAAT 2211
QY 2215 TTCACATCTATCCATGAACATCTGAACATGATGGTGGGGAAGTGAGAACTTACAATC 2274
DB 2212 TTCACATCTATCCATGAACATCTGAACATGATGGTGGGGAAGTGAGAACTTACAATC 2271
QY 2275 CAGGAAGGAAATGACGTATTTAAAGAGAAATACGTACACTACCGGGGACTTTTAAATGAG 2334
DB 2272 CAGGAAGGAAATGACGTATTTAAAGAGAAATACGTACACTACCGGGGACTTTTAAATGAG 2331
QY 2335 TGTATCCGACGTATTTATATCAAAAAATAGAGAGTCCGAAATTTAAAGCTTATATCTCGC 2394
DB 2332 TGTATCCGACGTATTTATATCAAAAAATAGAGAGTCCGAAATTTAAAGCTTATATCTCGC 2391
QY 2395 TACCAATTAAGGGGTATATTGAGATAGTCAAGATTTAGAGATATATTGATTCGTTAT 2454
DB 2392 TACCAATTAAGGGGTATATTGAGATAGTCAAGATTTAGAGATATATTGATTCGTTAT 2451
QY 2455 AATCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGTCCGCTTTCAGTT 2514
DB 2452 AATCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGTCCGCTTTCAGTT 2511
QY 2515 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGCACCACTTTTGAATGGAAT 2574
DB 2512 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGCACCACTTTTGAATGGAAT 2571
QY 2575 CCTGATCTAGATTGTTCTTCGAGAGATGGAGAAATGTGGCATCATTCCTCATCTTC 2634
DB 2572 CCTGATCTAGATTGTTCTTCGAGAGATGGAGAAATGTGGCATCATTCCTCATCTTC 2631
QY 2635 TCTTTGGATATTGATATTGGATGACAGACTTGCAATGAGATCTAGGCGTGTGGTGTGA 2694
DB 2632 TCTTTGGATATTGATATTGGATGACAGACTTGCAATGAGATCTAGGCGTGTGGTGTGA 2691
QY 2695 TTCAAGATTAAAGACGAGGAGGTCAATCAAGACTTAGGAACTTGGAAATTTATTGAAGAG 2754
DB 2692 TTCAAGATTAAAGACGAGGAGGTCAATCAAGACTTAGGAACTTGGAAATTTATTGAAGAG 2751
QY 2755 AAACCATTTTAGGAGAGCACTGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2814
DB 2752 AAACCATTTTAGGAGAGCACTGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2811
QY 2815 AAACCTGAAAACTCAATTTGAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2874
DB 2812 AAACCTGAAAACTCAATTTGAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2871
QY 2875 GATGCTTTATTGTTAGATTCTCAATATATATAGATTACAGCGGATACAAACATTTGGCATG 2934
DB 2872 GATGCTTTATTGTTAGATTCTCAATATATATAGATTACAGCGGATACAAACATTTGGCATG 2931
QY 2935 ATTCAATCGGAGAGATAAATCTGTTCAATTCGAATTCGAGAGGCTTATCTGTCAGAATTATCT 2994
DB 2932 ATTCAATCGGAGAGATAAATCTGTTCAATTCGAATTCGAGAGGCTTATCTGTCAGAATTATCT 2991
QY 2995 GTTATCCCGGGTGAATTCGGAATTTTGAAGAAATTTAGAGGTTCGCAATTTTCACTGCA 3054
DB 2992 GTTATCCCGGGTGAATTCGGAATTTTGAAGAAATTTAGAGGTTCGCAATTTTCACTGCA 3051
QY 3055 ATCTCCCTATACGATCGAGAGAAATGTCGTTAAAAATGTTGATTTTAAATAGGATTAGCA 3114

DB 3052 ATCTCCCTATACGATCGAGAGAAATGTCGTTAAAAATGTTGATTTTAAATAGGATTAGCA 3111
QY 3115 TGCTGGAATGTAAGAGGCACTGTAGATGTACACAGAGCCATCACCGTTCTGTCTCTGTT 3174
DB 3112 TGCTGGAATGTAAGAGGCACTGTAGATGTACACAGAGCCATCACCGTTCTGTCTCTGTT 3171
QY 3175 ATCCAGAAATCGGAGAGCAAGAGTGTCAAGCAGTTTCGCTCTGTCCGGGGCGTGGCTAT 3234
DB 3172 ATCCAGAAATCGGAGAGCAAGAGTGTCAAGCAGTTTCGCTCTGTCCGGGGCGTGGCTAT 3231
QY 3235 ATCTCCGTTGTACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAA 3294
DB 3232 ATCTCCGTTGTACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAA 3291
QY 3295 ATCCAGAAATACAGAGCAACTAAATTTTAAAAAATCTGTGAAGAGAGAGAGTGTATCCA 3354
DB 3292 ATCCAGAAATACAGAGCAACTAAATTTTAAAAAATCTGTGAAGAGAGAGAGTGTATCCA 3351
QY 3355 ACAGATACAGCAACGTTGTAATGATTTACTGCACACCAAGGTACAGCAGTATGTAATTC 3414
DB 3352 ACAGATACAGCAACGTTGTAATGATTTACTGCACACCAAGGTACAGCAGTATGTAATTC 3411
QY 3415 CGTAAATGCTGATATGAGGATGCAATATGAAGTTGATCTACAGCATCTGTTAAATTAACAA 3474
DB 3412 CGTAAATGCTGATATGAGGATGCAATATGAAGTTGATCTACAGCATCTGTTAAATTAACAA 3471
QY 3475 CCGACTTATGAAGAGAAACGTTATACAGATCTACAGAGAGATATCATTTGTAATATGAC 3534
DB 3472 CCGACTTATGAAGAGAAACGTTATACAGATCTACAGAGAGATATCATTTGTAATATGAC 3531
QY 3535 AGAGGATGTTGAATTTATCCACCACTACCGCTGTTATATATGACAAAAAGAAATAGAAATAC 3594
DB 3532 AGAGGATGTTGAATTTATCCACGATACCGCTGTTATATATGACAAAAAGAAATAGAAATAC 3591
QY 3595 TTCCAGAAACCGATAAGGTATGATGAGATTGGAGAAACGGAAGGAAAGTTTATTGTA 3654
DB 3592 TTCCAGAAACCGATAAGGTATGATGAGATTGGAGAAACGGAAGGAAAGTTTATTGTA 3651
QY 3655 GACAGCGTGGAAATTTACTCTTATCGAGGAATAG 3687
DB 3652 GACAGCGTGGAAATTTACTCTTATCGAGGAATAG 3684

RESULT 4

US-09-661-322A-62
; Sequence 62, Application US/09661322A
; Patent No. 6593293

GENERAL INFORMATION:

; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Ruper, Mark J.

; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Compos

; FILE REFERENCE: MISC0201

; CURRENT APPLICATION NUMBER: US/09/661,322A

; CURRENT FILING DATE: 2000-09-13

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 62

; LENGTH: 3684

; TYPE: DNA

; ORGANISM: Bacillus thuringiensis

US-09-661-322A-62

Query Match 88.4%; Score 3260.2; DB 4; Length 3684;

Best Local Similarity 93.3%; Pred. No. 0;

Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

QY 1 TTGACTTCAATAGGAAAAATGAGAATTAATAATGCTTTTATCGATCCAGCTGTA 60

Db 1 TTGACTTCAAAATAGGAAAAATGGAATGAAATTTATAAATGCTTTATGCAATCCAGCTGTA 60
Qy 61 TCGAATCATTTCCACACAAATGGAATCTATCAACAGATGCTCGTATTGAGGATTTCTTTGGT 120
Db 61 TCGAATCATTTCCGACAAATGAATCTATCAACGATGCTCGTATTGAGGATAGCTTGGT 120
Qy 121 ATAGCCGAGGGGAATAAATCAATCACTGTTGAGCGCATCAACAGTCAAAACGGGTATT 180
Db 121 ATAGCCGAGGGGAACAAATATCGATCCATTTGTTAGCGCATCAACAGTCAAAACGGGTATT 180
Qy 181 AACATAGCTGGTAGAATACATAGTGTATTAGGGTACCGTTTGGCGACAAATAGCTAGT 240
Db 181 AACATAGCTGGTAGAATACATAGTGTATTAGGGTACCGTTTGGCGACAAATAGCTAGT 240
Qy 241 TTTTATAGTTTCTTGTGTGGTAATATATGCCCCCGCGGAGAGATCAGTGGGAAATTTTC 300
Db 241 TTTTATAGTTTCTTGTGTGGTAATATATGCCCCCGCGGAGAGATCCTTTGGGAAATTTTC 300
Qy 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATTAACGAAATATCTAGGGATACGGCA 360
Db 301 CTAGAACATGTCGAACAACTTATAAAGCAACAAAGTAAACGAAATATCTAGGGATACGGCT 360
Qy 361 CTTGCTCGATTACAAGGTTTAGAGATTCCTTTAGAGCCTTATCAACAGTCACTTTGAAGAT 420
Db 361 CTTGCTCGATTACAAGGTTTAGGAAATTCCTTTAGAGCCTTATCAACAGTCACTTTGAAGAT 420
Qy 421 TGGCTAGAAAAACCGTGATGATGCAAGAAACGAGAAGTGTCTTTATACCCAAATATATAGCC 480
Db 421 TGGCTAGAAAAACCGTGATGATGCAAGAAACGAGAAGTGTCTTTATACCCAAATATATAGCC 480
Qy 481 TTAGAACTTGATTTCTTAATGCGATGCCGCTTTTCCAAATTAGAAACCAAGAGTTCCA 540
Db 481 TTAGAACTTGATTTCTTAATGCGATGCCGCTTTTCCAAATTAGAAACCAAGAGTTCCA 540
Qy 541 TTATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTAATTTAGAGATGCTCT 600
Db 541 TTATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTAATTTAGAGATGCTCT 600
Qy 601 CTTTTTGGTAGTAATTTGGCTTTACATCGCAGGAAATCAACGTTTATATAGAGCGCAA 660
Db 601 CTTTTTGGTAGTAATTTGGCTTTACATCCCAAGAAATCAACGTTTATATAGAGCGCAA 660
Qy 661 GTGAAACAAACGAGAGATTTCCGACTATTTGCGTAGAATGGTATATACAGTCTAAAT 720
Db 661 GTGAAACAAACGAGAGATTTCTGATTTATGCGCAAGATGGTATATACGAGTTTAAAT 720
Qy 721 AGCTTGAGAGGACAAATGCCGCAAGTTGGGTGGTGTATTAATCAATTCCTGAGAGTCTA 780
Db 721 AATTTGAGAGGACAAATGCTGAAAGTTGGTTGCGATATTAATCAATTCCTGAGAGCTTA 780
Qy 781 ACGTTAGGGGTATTAGATCTAGTGCGACTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
Db 781 ACGCTAGGAGTATTAGATCTAGTGCGACTATTTCCCAAGCTATGACACGCGTGTATTATCCA 840
Qy 841 ATAATACAGTGCTCAGTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
Db 841 ATGAATACAGTGCTCAATTAACAGAGAAATTTATACAGATCCAAATTTGGAGAACAAAT 900
Qy 901 GTAAAT-----ATGGCAAGTATGAATTTGGTATTAATTAATATGCACTTCCTGTTTCGCT 954
Db 901 GCACCTTCAGGATTTTGCAAGTACGAATTTGGTTTAAATAATAATGCAACCATCTGTTTCGCT 960
Qy 955 ATAGAGACTGGGTTTATCGAAGCCGCATCTACTTGATTTTCTAGAACACTTTACAAAT 1014
Db 961 ATAGAGGCTCGCGTTATTAGGCTCCGCATCTACTTTGATTTTTCAGAACACTTTACAAAT 1020
Qy 1015 TTTAGCACTTCATACAGATGAGTGTCTACTAGGCATATGACTTTACTGGCGGGGCGACACA 1074
Db 1021 TTCAGGATTTAAGTCGATGGAGTAAATCTCAATATATGAATTAATCTGGGTGGACATAGA 1080
Qy 1075 ATTTCAATCTCGCCCAATAGGAGCGGATTTAAATACCTCAACGATGGGTCTACCAATACT 1134
Db 1081 CTTGAATCGCGAAACAAATAGGGGGGTCAITTAAGTACCTGGACACACGGAATACCAATACT 1140

Qy 1135 TCTATTAACTCTGTGAAGATTATCATTTCTCTCGAGACGTTATATTGGACTGAATCATAT 1194
Db 1141 TCTATTAACTCTGTGAACATTACAGTTTACATCTCGAGACGTTTATATGAACAGAAATCATTT 1200
Qy 1195 GCAGGAGTGCTTCTATGGGGAATTTACCTTTGAACCTTATTATGAGTGTCCCTACTGTTAGA 1254
Db 1201 GCAGGATAAATAT-----ACTTCTAACTACTCTCTGTAATGGAGTACCTTTGGGTAGA 1254
Qy 1255 TTTAATTTTAGGAACCTCTGAATATCTTTTGAAGAGGTACTGCTAACTACTAGTCAACCC 1314
Db 1255 TTTAATTTAGGAATATCCCTGAAATTCCTT-----AGAGGTAGCCTTCTCTACTATATAGG 1311
Qy 1315 TATGAGTCACCTGGGCTTCAATTAAGATTTACAGAACTGAATTTACCAACAGAAACAAACA 1374
Db 1312 TATACTGGAGTGGGGAACAACACTATTTGATTTCAAGAACTGAAATTTACCAACAGAAACAAACA 1371
Qy 1375 GAAACACCAAAATATGAATCATATAGTCATAGGTTATCTCAATAGGGCTCATTTTCAACA 1434
Db 1372 GAAACACCAAAATATGAATCTTACAGTCATAGATTATCTAATAATAAGACTAATATACGGA 1431
Qy 1435 TCTAGGGTGCAATGATACAGTATATCTTGGACGACCGTAGTGCAGATCGTACAAATACC 1494
Db 1432 AACACTTTGAGAGCACCAAGTATATCTTGGAGCACCGTAGTGCAGATCGTACAAATACC 1491
Qy 1495 ATTAGTTCAGATAGCATAAACAAATACCATTTGGTAAATCAITTCACCTTTAAATTCAGGT 1554
Db 1492 ATTAGTTCAGATAGCATAAACAAATACCATTTGGTAAATCAITTCACCTTTAAATTCAGGT 1551
Qy 1555 ACCTCTGAGTCAGTGGCCCGAGGATTTACAGAGGGGATATAATCCGAACCTTAACGTTAAT 1614
Db 1552 ACCTCTGAGTCAGTGGCCCGAGGATTTACAGAGGGGATATAATCCGAACCTTAACGTTAAT 1611
Qy 1615 GGTAGTGTACTAAGTATGGGTCTTAATTTAATAATACATCATTTACACGGGTATCGCGTG 1674
Db 1612 GGTAGTGTACTAAGTATGGGTCTTAATTTAATAATACATCATNTACAGCGGTATCGCGTG 1671
Qy 1675 AGAGTTTCGTTATGCTGCTTCTCAAAACAAATGGTCTCGAGGTAACCTGTCCGAGGGAGTACT 1734
Db 1672 AGAGTTTCGTTATGCTGCTTCTCAAAACAAATGGTCTCGAGGTAACCTGTCCGAGGGAGTACT 1731
Qy 1735 ACTTTTGTATCAAGGATTCCTTAGTACTAGTGCAGAAATAGTCTTTGACATCTCAATCA 1794
Db 1732 ACTTTTGTATCAAGGATTCCTTAGTACTAGTGCAGAAATAGTCTTTGACATCTCAATCA 1791
Qy 1795 TTTAGATTTGCAGAAATTTCTCTAGTATTAGTGCATCTGGCAGTCAAACTGCTCGAATA 1854
Db 1792 TTTAGATTTGCAGAAATTTCTCTAGTATTAGTGCATCTGGCAGTCAAACTGCTCGAATA 1851
Qy 1855 AGTATAAGTAAATATGACAGGTAGACAAACGTTTCACTTTTGATAAAAATTGAAATTCATTCCA 1914
Db 1852 AGTATAAGTAAATATGACAGGTAGACAAACGTTTCACTTTTGATAAAAATTGAAATTCATTCCA 1911
Qy 1915 ATTACTGCAACCTTTGAGCAGAAATACGATTTAGAAAGGCGCCAGAGCGGTGAATGCT 1974
Db 1912 ATTACTGCAACCTTTGAGCAGAAATATGATTTAGAAAGGCGCCAGAGCGGTGAATGCT 1971
Qy 1975 CTGTTTACTTAATAACGAATCCAAAGAGATTTGAAAAACAGATGTGACAGATTATCATATTGAT 2034
Db 1972 CTGTTTACTTAATAACGAATCCAAAGAGTTTGAACACAGGTGTGACAGATTATCATATTGAT 2031
Qy 2035 CAAAGTATCCAAATTTAGTGGCGTTTATCGGATGAATTTCTGTTAGATGAAGAGAGAA 2094
Db 2032 GAAAGTATCCAAATTTAGTGGCGTTTATCGGATGAATTTCTGCTTGGATGAAGAGAGAA 2091
Qy 2095 TTTACTTGAGAAAGTGAATATCGAAACGACTCAGTGTGAAAGAACTTTACTTCCAAGAT 2154
Db 2092 TTTACTTGAGAAAGTGAATATCGAAACGACTCAGTGTGAAAGAACTTTACTTCCAAGAT 2151
Qy 2155 CCAAACTTCACTCCCATCAATAAGCAACCAAGCTTCAATCTCTAATATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCACTCCCATCAATAAGCAACCAAGCTTCAATCTCTAATATGAGCAATCGAAT 2211

2215 TTCAATCTATCCATGAACAACTCTGAACATGGATGGTGGGAGTGAAGAACATTCATCAATC 2274
2212 TTCAATCTATCCATGAACAACTCTGAACATGGATGGTGGGAGTGAAGAACATTCATCAATC 2271
2275 CAGGAGGAAATGACGTATTTAAAGAGATTAAGTCAACACTACCGGGAGCTTTTAATGAG 2334
2272 CAGGAGGAAATGACGTATTTAAAGAGAAATTAAGTCAACACTACCGGGAGCTTTTAATGAG 2331
2335 TGTTATCCGACGTATTTATATCAAAAAATAGGAGAGTCCGAATTAAGCTTATATACTGCG 2394
2332 TGTTATCCGACGTATTTATATCAAAAAATAGGAGAGTCCGAATTAAGCTTATATACTGCG 2391
2395 TACCAATTAAGAGGTATATTGAAGATCAAGATTTAGAGATATATTTGATTCGTTAT 2454
2392 TACCAATTAAGTGGCTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAC 2451
2455 AATGCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTCAGTT 2514
2452 AATGCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTCAGTT 2511
2515 GAAAGCCCAATCGGAGGTGCGAGAACCGAATCGATCGCACACATTTTGAATGGAAT 2574
2512 GAAAGCCCAATCGGAGGTGCGAGAACCGAATCGATCGCACACATTTTGAATGGAAT 2571
2575 CCTGATCTAGATTTGTTCTGTCAGAGATGGAGAAAAATGTGCGCATCATTTCCCATCATTTTC 2634
2572 CCTGATCTAGATTTGTTCTGTCAGAGATGGAGAAAAATGTGCGCATCATTTCCCATCATTTTC 2631
2635 TCTTTGGATATTGATATTGGATGTCAGACACTTTGCHATGAGAACTTAGGCGTGTGGTGGTA 2694
2632 TCTTTGGATATTGATATTGGATGTCAGACACTTTGCHATGAGAACTTAGGCGTGTGGTGGTA 2691
2695 TTCAAGATTTAAGACCGAGGAGTCTATGCAAGACTTAGGAACTCTGGAATTTATTGAAGAG 2754
2692 TTCAAGATTTAAGACCGAGGAGTCTATGCAAGACTTAGGAACTCTGGAATTTATTGAAGAG 2751
2755 AAACCATTTATTAGGAGAGACTGTCTCGTGTGAAGAGAGCAGAGAAAAATGGAGAGAC 2814
2752 AAACCATTTATTAGGAGAGACTGTCTCGTGTGAAGAGAGCAGAGAAAAATGGAGAGAC 2811
2815 AAACCTGAAACACTCAATTTGAAACAAACCGATATATACAGAGGCAAAAGAGCTGTG 2874
2812 AAACCTGAAACACTCAATTTGAAACAAACCGATATATACAGAGGCAAAAGAGCTGTG 2871
2875 GATGCTTTATTGTAGATTCTCAATATATAGATTACAAGCGGATACAAACATTCGSCATG 2934
2872 GATGCTTTATTGTAGATTCTCAATATATAGATTACAAGCGGATACAAACATTCGSCATG 2931
2935 ATTCAATCGGCGAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTGTCAGAATTATCT 2994
2932 ATTCAATCGGCGAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTTTCAGAATTATCT 2991
2995 GTTATCCCGGTGTAAATTCGGAAATTTTGAAGAAATTAAGAGTTCGATTTATCACTGCA 3054
2992 GTTATCCCGGTGTAAATTCGGAAATTTTGAAGAAATTAAGAGTTCGATTTATCACTGCA 3051
3055 ATCTCCCTATACGATCGGAGAAATCGTTAAATGTTGATTTTAAATGGAATTAGCA 3114
3052 ATCTCCCTATACGATCGGAGAAATCGTTTAAATGTTGATTTTAAATGGAATTAGCA 3111
3115 TGCTGGAATGTAAAAAGGCGATGTAGATGTACAACAGAGCCATCACCGTTCTGCTCTGTT 3174
3112 TGCTGGAATGTAAAAAGGCGATGTAGATGTACAACAGAGCCATCACCGTTCTGCTCTGTT 3171
3175 ATCCAGAAATGGGAAGCAGAGTGTCAAGCAGTTCCGGTGTGTCGGGGCGTGGCTAT 3234
3172 ATCCAGAAATGGGAAGCAGAGTGTCAAGCAGTTCGCTGTGTCGGGGCGTGGCTAT 3231
3235 ATCTCTCGGTGTACAGCGGTACAAAGAGGATATGAGAGGTTGTTAAACGATCCATGAA 3294
3232 ATCTCTCGGTGTACAGCGGTACAAAGAGGATATGAGAGGTTGTTAAACGATCCATGAA 3291
3295 ATCGAGAACAAATACAGAGCAACTAAAAATTTAAAAAATCTGTGAAGAGAGGAGTGTATCCA 3354

3292 ATCGAGAACAAATACAGAGCAACTAAAAATTTAAAAAATCTGTGAAGAGAGGAGTGTATCCA 3351
3355 ACGGATACAGGAACGTGTAAATGATTATATCTGACACACCAAGGTACAGCAGTATGTAATCC 3414
3352 ACGGATACAGGAACGTGTAAATGATTATATCTGACACCAAGGTACAGCAGTATGTAATCC 3411
3415 CGTAATGCTGGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTTAAATTACAAA 3474
3412 CGTAATGCTGGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTTAAATTACAAA 3471
3475 CCGACTTATGAAGAGAAACCGTATATACAGATGTACGAAGAGATAATCATTTGTAATATGAC 3534
3472 CCGACTTATGAAGAGAAACCGTATATACAGATGTACGAAGAGATAATCATTTGTAATATGAC 3531
3535 AGAGGGTATGTGAATTTATCCACCTACCGCTGTTTATATGACAAAAGAAATTAGAAATAC 3594
3532 AGAGGGTATGTGAATTTATCCACCTACCGCTGTTTATATGACAAAAGAAATTAGAAATAC 3591
3595 TTCCAGAGAACCGGATTAAGGTATGAGATTGAGATTGGAGAAAACGAAAGGAGGTTTATTGTA 3654
3592 TTCCAGAGAACCGGATTAAGGTATGAGATTGAGATTGGAGAAAACGAAAGGAGGTTTATTGTA 3651
3655 GACAGCGTGGAAATTAATCTCTTATGGAGGAATAG 3687
3652 GACAGCGTGGAAATTAATCTCTTATGGAGGAATAG 3684

RESULT 5

US-08-377-690-1
; Sequence 1, Application US/08377690
; Patent No. 5628995
; GENERAL INFORMATION:
; APPLICANT: PEPEROEN, Marnix
; APPLICANT: JANSSENS, Stefan
; APPLICANT: DENOLF, Peter
; TITLE OF INVENTION: CONTROL OF OSTRINIA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: The George Mason Bldg., Washington & Prince
; STREET: Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,690
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/164,781
; FILING DATE:
; APPLICATION NUMBER: US 07/938,362
; FILING DATE: 31-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feuty, Sharon B
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-039
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4074 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: entomocidus HD 110
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 186..3872
; OTHER INFORMATION: /note= "PROPERTIES: CryIB is toxic to
; OTHER INFORMATION: Ostrinea nubilalis (among others)"
; US-08-377-690-1

Query Match      85.0%; Score 3133.6; DB 1; Length 4074;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 299; Indels 30; Gaps 3;

QY 1 TTGACTTCAAAATAGGAAAAATGAGATGAAATTAATAATGCTTTATCGATTCGAGCTGTA 60
DB 186 TTGACTTCAAAATAGGAAAAATGAGATGAAATTAATAATGCTTTATCGATTCGAGCTGTA 230
QY 61 TCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTGAGGATTTCTTTGTGT 120
DB 231 TCGAATCATTTCCGACAAATGGATCTATTACCAGATGCTCGTATTGAGGATGCTTTGTGT 290
QY 121 ATAGCCGAGGGGAATTAATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT 180
DB 291 ATAGCCGAGGGGAATTAATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT 350
QY 181 AACATAGCTGGTGAATACATAGGTGTATTAGGCGTACCGTTTCTGGACAAATAGTAGT 240
DB 351 AACATAGCTGGTGAATACATAGGCGTATTGGGCGTACCGTTTCTGGACAAATAGTAGT 410
QY 241 TTTTATAGTTTCTCTGTGGTGAATATTGCGCCGCGGAGAGATCAGTGGGAAATTTTC 300
DB 411 TTTTATAGTTTCTCTGTGGTGAATATTGCGCCGCGGAGAGATCAGTGGGAAATTTTC 470
QY 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATAGCTAGGAATACGGCA 360
DB 471 CTAGAACATGTCGAACAACTTATAAATCAACAAATAGCTAGGAATACGGCT 530
QY 361 CTTGCTCGATTACAGGTTTAGAGATTCCTTTAGAGCTTATCAACAGTCACTTGAAGAT 420
DB 531 CTTGCTCGATTACAGGTTTAGAGATTCCTTTAGAGCTTATCAACAGTCACTTGAAGAT 590
QY 421 TGCGTAGAAACCGTGATGATCAAGAACGAGAGTGTCTTTATACCCAAATATAGCC 480
DB 591 TGCGTAGAAACCGTGATGATCAAGAACGAGAGTGTCTTTATACCCAAATATAGCT 650
QY 481 TTAGAACTTGATTTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAGTTCCA 540
DB 651 TTAGAACTTGATTTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAGTTCCA 710
QY 541 TTATTAAATGATATGCTCAAGCTGCAAAATTTACCTATTATTATTGAGAGATGCTCT 600
DB 711 TTATTAAATGATATGCTCAAGCTGCAAAATTTACCTATTATTATTGAGAGATGCTCT 770
QY 601 CTTTTTTGGTAGTGAATTTGGGCTTACATCCGAGGAATTCAGCTTTATTATGAGCCCA 660
DB 771 CTTTTTTGGTAGTGAATTTGGGCTTACATCCGAGGAATTCAGCTTTATTATGAGCCCA 830
QY 661 GTGGAACAAACGAGAGATTATTCGACTATTGGGTAGAAATGGTATATACAGGTCTAAAT 720
DB 831 GTGGAACAAACGAGAGATTATTCGACTATTGGGTAGAAATGGTATATACAGGTCTAAAT 890
QY 721 AGCTTTAGAGGGAACAAATGCGGAAATTTGGGTGGTGTATTATCAATTTCCGTAGAGATCTA 780
DB 891 AGCTTTAGAGGGAACAAATGCGGAAATTTGGGTGGTGTATTATCAATTTCCGTAGAGATCTA 950
QY 781 ACGTTAGGGGTATTAGATCTAGTGGCACTATCCCAAGCTTATGACACTCCGACTTATCCA 840
DB 951 ACGTTAGGGGTATTAGATCTAGTGGCACTATCCCAAGCTTATGACACTCCGACTTATCCA 1010
QY 841 ATAAATACGAGTGTCTAGTTTAAACAAAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
DB 900 ATAAATACGAGTGTCTAGTTTAAACAAAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 960
QY 960 GTAAATATGGCAAGTATGAATTTGGTATATATATATATGACACCTTCGTTTCCGCTATAGAG 1020
DB 1071 GTAAATATGGCAAGTATGAATTTGGTATATATATATATGACACCTTCGTTTCCGCTATAGAG 1130
QY 961 ACTGGGGTTATCCGAAAGCCCGCATCTACTTGTATTTTGTAGAACAACTTACAAATTTTTC 1020
DB 1131 GCTGGGGTTATCCGAAAGCCCGCATCTACTTGTATTTTGTAGAACAACTTACAAATTTTTC 1190
QY 1021 ACTTCATACGATGAGTGTCTACTAGGCAATATGACTTACTGGGGGGGACACAATTTCAA 1080
DB 1191 GCTTCATACGATGAGTGTCTACTAGGCAATATGACTTACTGGGGGGGACACAATTTCAA 1250
QY 1081 TCTCGGCCAATAGGAGGCGGATTAATACCTCAACGATGGGTCTACCAATCTCTTATT 1140
DB 1251 TCTCGGCCAATAGGAGGCGGATTAATACCTCAACGATGGGTCTACCAATCTCTTATT 1310
QY 1141 AATCTCTGAAGATTATCATTTCTCTCGAGACGTATATTTGGACTGAAATCATATGCAGGA 1200
DB 1311 AATCTCTGAAGATTATCATTTCTCTCGAGACGTATATTTAGGACTGAAATCATATGCAGGA 1370
QY 1201 GTGCTTCTATGGGAAATTTTACCTTGAACTTATCATGTTGTCCTACTCTGTTAGATTAAAT 1260
DB 1371 GTGCTTCTATGGGAAATTTTACCTTGAACTTATCATGTTGTCCTACTCTGTTAGATTAAAT 1430
QY 1261 TTTAGGAAACCTTCAGATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATAG 1320
DB 1431 TTTAGGAAACCTTCAGATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATAG 1490
QY 1321 TCACCTGGGCTTCAATTAAGAAATTCAGAAATCTGAATTAACCAAGAAACAAACAGAACGA 1380
DB 1491 TCACCTGGGCTTCAATTAAGAAATTCAGAAATCTGAATTAACCAAGAAACAAACAGAACGA 1550
QY 1381 CCAATTTATGAATCATATAGTCAATAGTTATCTCATATAGGGGTCAATTCACAACTACCTAG 1440
DB 1551 CCAATTTATGAATCATATAGTCAATAGTTATCTCATATAGGGGTCAATTCACAACTACCTAG 1610
QY 1441 GTGCAATGACAGTATATTTCTTGGAGCGCATCTGAGTGCAGATCGTACAAATACCAATTAGT 1500
DB 1611 GTGCAATGACAGTATATTTCTTGGAGCGCATCTGAGTGCAGATCGTACAAATACCAATTAGT 1570
QY 1501 TCAGATAGCATAAACCAAAATACCATTTGGTAAATCAATTCACCTTAATTCAGGTACCTCT 1560
DB 1671 CCAATTAGAATCAACCCAAATCCCAATGGTAAAGCATCCGAACTTCTCTCAAGGTACCACT 1730
QY 1561 GTAGTCAGTGGCCAGGATTTACAGAGGGGATTAATTCGAACTAACGTTTAAATGTTAGT 1620
DB 1731 GTTGTTAGAGGACCAAGATTTTACTGTTGGGGATTAATTCGAAAGAACGAATACTGTTGGA 1790
QY 1621 GTACTAAGTATGGGTCTTAATTTTAAATATACATCATTACAGCGGTATCGCGTGAGAGTT 1680
DB 1791 TTTGGACCGATAGAGTAACTGTTAACGGACCAATTAACCAAGATATCGTATAGATTC 1850
QY 1681 CGTTATGCTGCTTCTCAAAATGGTCTCGAGGGTAACTGTTCGAGGGGAGTACTACTTTT 1740
DB 1851 CGCTATGCTTCACTGTAGATTTTGTATTTTGTATCACGTGGAGGTACTACTGTAAAT 1910
QY 1741 GATCAAGGATTCCTTAGTACTAGTGCAGAAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
DB 1911 AATTTTAGATTTCTACGTACAATGAACAGTGGAGACGAATAAATACGGAATTTTGTG 1970
QY 1801 TTTGCAAGATTTCTCTAGGTATTAGTGCATCTGGCAGTCAA--ACTGCTGGAATAAGT 1857
DB 1971 AGACGTGCTTTTACTACCTTTTACTTTTACAAATTCAGATATTAATTCGAACTCT 2030
QY 1858 ATAAGTAATAATGCGAGTAGACAAACGTTTCACTTTTGATAAAATTTGAATTCATTCCAAAT 1917
DB 2031 ATTCAAGCCCTTAGTGGAAATGGGAAGTGTATATAGATAAAATTTGAAATTTATCCAGTT 2090
QY 1918 ACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGCGGACAGCGGTGAATGCTCTG 1977
DB 2091 ACTGCAACCTTCGAAGCAGAAATATGATTTTAGAAAGAGCGCAAGAGCGGTGAATGCTCTG 2150
```

```
DB 1011 ATAAATACGAGTGTCTCAGTTAAACAGAGAAAGTTTATACAGACGCAATTTGGAGCAACAGG 1070
QY 901 GTAAATATGGCAAGTATGAATTTGGTATATATATATATGACACCTTCGTTTCCGCTATAGAG 960
DB 1071 GTAAATATGGCAAGTATGAATTTGGTATATATATATATGACACCTTCGTTTCCGCTATAGAG 1130
QY 961 ACTGGGGTTATCCGAAAGCCCGCATCTACTTGTATTTTGTAGAACAACTTACAAATTTTTC 1020
DB 1131 GCTGGGGTTATCCGAAAGCCCGCATCTACTTGTATTTTGTAGAACAACTTACAAATTTTTC 1190
QY 1021 ACTTCATACGATGAGTGTCTACTAGGCAATATGACTTACTGGGGGGGACACAATTTCAA 1080
DB 1191 GCTTCATACGATGAGTGTCTACTAGGCAATATGACTTACTGGGGGGGACACAATTTCAA 1250
QY 1081 TCTCGGCCAATAGGAGGCGGATTAATACCTCAACGATGGGTCTACCAATCTCTTATT 1140
DB 1251 TCTCGGCCAATAGGAGGCGGATTAATACCTCAACGATGGGTCTACCAATCTCTTATT 1310
QY 1141 AATCTCTGAAGATTATCATTTCTCTCGAGACGTATATTTGGACTGAAATCATATGCAGGA 1200
DB 1311 AATCTCTGAAGATTATCATTTCTCTCGAGACGTATATTTAGGACTGAAATCATATGCAGGA 1370
QY 1201 GTGCTTCTATGGGAAATTTTACCTTGAACTTATCATGTTGTCCTACTCTGTTAGATTAAAT 1260
DB 1371 GTGCTTCTATGGGAAATTTTACCTTGAACTTATCATGTTGTCCTACTCTGTTAGATTAAAT 1430
QY 1261 TTTAGGAAACCTTCAGATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATAG 1320
DB 1431 TTTAGGAAACCTTCAGATACCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATAG 1490
QY 1321 TCACCTGGGCTTCAATTAAGAAATTCAGAAATCTGAATTAACCAAGAAACAAACAGAACGA 1380
DB 1491 TCACCTGGGCTTCAATTAAGAAATTCAGAAATCTGAATTAACCAAGAAACAAACAGAACGA 1550
QY 1381 CCAATTTATGAATCATATAGTCAATAGTTATCTCATATAGGGGTCAATTCACAACTACCTAG 1440
DB 1551 CCAATTTATGAATCATATAGTCAATAGTTATCTCATATAGGGGTCAATTCACAACTACCTAG 1610
QY 1441 GTGCAATGACAGTATATTTCTTGGAGCGCATCTGAGTGCAGATCGTACAAATACCAATTAGT 1500
DB 1611 GTGCAATGACAGTATATTTCTTGGAGCGCATCTGAGTGCAGATCGTACAAATACCAATTAGT 1570
QY 1501 TCAGATAGCATAAACCAAAATACCATTTGGTAAATCAATTCACCTTAATTCAGGTACCTCT 1560
DB 1671 CCAATTAGAATCAACCCAAATCCCAATGGTAAAGCATCCGAACTTCTCTCAAGGTACCACT 1730
QY 1561 GTAGTCAGTGGCCAGGATTTACAGAGGGGATTAATTCGAACTAACGTTTAAATGTTAGT 1620
DB 1731 GTTGTTAGAGGACCAAGATTTTACTGTTGGGGATTAATTCGAAAGAACGAATACTGTTGGA 1790
QY 1621 GTACTAAGTATGGGTCTTAATTTTAAATATACATCATTACAGCGGTATCGCGTGAGAGTT 1680
DB 1791 TTTGGACCGATAGAGTAACTGTTAACGGACCAATTAACCAAGATATCGTATAGATTC 1850
QY 1681 CGTTATGCTGCTTCTCAAAATGGTCTCGAGGGTAACTGTTCGAGGGGAGTACTACTTTT 1740
DB 1851 CGCTATGCTTCACTGTAGATTTTGTATTTTGTATCACGTGGAGGTACTACTGTAAAT 1910
QY 1741 GATCAAGGATTCCTTAGTACTAGTGCAGAAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
DB 1911 AATTTTAGATTTCTACGTACAATGAACAGTGGAGACGAATAAATACGGAATTTTGTG 1970
QY 1801 TTTGCAAGATTTCTCTAGGTATTAGTGCATCTGGCAGTCAA--ACTGCTGGAATAAGT 1857
DB 1971 AGACGTGCTTTTACTACCTTTTACTTTTACAAATTCAGATATTAATTCGAACTCT 2030
QY 1858 ATAAGTAATAATGCGAGTAGACAAACGTTTCACTTTTGATAAAATTTGAATTCATTCCAAAT 1917
DB 2031 ATTCAAGCCCTTAGTGGAAATGGGAAGTGTATATAGATAAAATTTGAAATTTATCCAGTT 2090
QY 1918 ACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGCGGACAGCGGTGAATGCTCTG 1977
DB 2091 ACTGCAACCTTCGAAGCAGAAATATGATTTTAGAAAGAGCGCAAGAGCGGTGAATGCTCTG 2150
```

Qy	1978	TTTACTAATACGAATCCAGAAGATTGAAAACAGATGTGACAGATTATCATATTGATCAA	2037
Db	2151	TTTACTAATACGAATCCAGAAGATTGAAAACAGATGTGACAGATTATCATATTGATCAA	2210
Qy	2038	GTATCCAAATTTAGTGGCGTGTATCCGATGGAATTCCTGCTTAGATGAAAAGACAGAAATTA	2097
Db	2211	GTATCCAAATTTAGTGGCGTGTATCCGATGGAATTCCTGCTTAGATGAAAAGACAGAAATTA	2270
Qy	2098	CTTGAGAAAGTGAAATATGCGAAACGACTCAGTGAATGAAAAGAACTTTACTCCAGATCCA	2157
Db	2271	CTTGAGAAAGTGAAATATGCGAAACGACTCAGTGAATGAAAAGAACTTTACTCCAGATCCA	2330
Qy	2158	AACCTTCATCTCAATTAAGAAACGACATTCATATCTACTAATAGCAATCGAATTC	2217
Db	2331	AACCTTCATCTCAATTAAGAAACGACATTCATATCTACTAATAGCAATCGAATTC	2390
Qy	2218	ACATCTATCCATGAACAATCTGAACATGGATGGTGGGGAAGTGGAGACATTACAATCCAG	2277
Db	2391	ACATCTATCCATGAACAATCTGAACATGGATGGTGGGGAAGTGGAGACATTACAATCCAG	2450
Qy	2278	GAAGAAATGACGTATTTAAAGAAATTACGTACACTACCGGGGACTTTTAATGAGTGT	2337
Db	2451	GAAGAAATGACGTATTTAAAGAAATTACGTACACTACCGGGTACTTTTAATGAGTGT	2510
Qy	2338	TATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATATCTCCTAC	2397
Db	2511	TATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATATCTCCTAC	2570
Qy	2398	CAATTAGAGGGTATATGGAAGATAGTCAAGATTAGAGATATATTTGATTCTGTTAAT	2457
Db	2571	CAATTAGAGGGTATATGGAAGATAGTCAAGATTAGAGATATATTTGATTCTGTTAAT	2630
Qy	2458	GCGAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGGCCGCTTTTCAGTTGAA	2517
Db	2631	GCGAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCCTATGGCCGCTTTTCAGTTGAA	2690
Qy	2518	AGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGGCACCAATTTTGAATCGAATCCT	2577
Db	2691	AGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGGCACCAATTTTGAATCGAATCCT	2750
Qy	2578	GATCTAGATTGTTCTCGCAGAGATGGAGAAAAATGTGCGGATCATTTCCCATCATTTCTCT	2637
Db	2751	GATCTAGATTGTTCTCGCAGAGATGGAGAAAAATGTGCGGATCATTTCCCATCATTTCTCT	2810
Qy	2638	TTGGATATTCGATATTTGGATGCGACAGCTTCGATCAGAACTCAGCGGTGGGTGGTATTC	2697
Db	2811	TTGGATATTCGATATTTGGATGCGACAGCTTCGATCAGAACTCAGCGGTGGGTGGTATTC	2870
Qy	2698	AAGATTAAAGACGGAGAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAA	2757
Db	2871	AAGATTAAAGACGGAGAGGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAA	2930
Qy	2758	CCATTATTAGAGAAGCACCTGTCTGTGTGAAGAGAGCAGAGAAAAAATGGAGAGACAA	2817
Db	2931	CCATTATTAGAGAAGCACCTGTCTGTGTGAAGAGAGCAGAGAAAAAATGGAGAGACAA	2990
Qy	2818	CGTGAAAACTACAAATTTGAAAACAAAACGAGTATATACAGAGCCAAAAGAGCTGTGGAT	2877
Db	2991	CGTGAAAACTACAAATTTGAAAACAAAACGAGTATATACAGAGCCAAAAGAGCTGTGGAT	3050
Qy	2878	GCTTTATTTGTAGATTCTCAATATAATAGATTACAAGCCGATACAAACATTTGCGCATGATT	2937
Db	3051	GCTTTATTTGTAGATTCTCAATATGATAGATTACAAGCCGATACAAACATTCGCGCATGATT	3110
Qy	2938	CATCGCCGAGATAAACTTTGTTTCATCGAAATTCGAGAGCTTTATCTGTGAGAAATATCTGTT	2997
Db	3111	CATCGCCGAGATAAACTTTGTTTCATCGAAATTCGAGAGCTTTATCTTTGAGAAATTTACCTGTT	3170
Qy	2998	ATCCCGGGTGTAATGCGGAAAATTTTGAAGAAATTAAGAGGTCGCAATTAATCACTGCAATC	3057
Db	3171	ATCCCGGGTGTAATGCGGAAAATTTTGAAGAAATTAAGAGGTCACATTAATCACTGCAATC	3230

3058	Qy	TCCTATACAGTCGGAGAAATGTCGTTTAAAAATGGGTGATTTTAAATATGATAGCATGC	3111
3231	Db	TCCTTATACAGTCGGAGAAATGTCGTTTAAAAATGGGTGATTTTAAATATGATAGCATGT	3290
3118	Qy	TGCAATGTAAGAGGGCATGTAGATGTACAACAGAGCCATCACCGTTCGTGCTTGTGTTATC	3177
3291	Db	TGGNAATGTAAGAGGGCATGTAGATGTACAACAGAGCCATCATCGTTCGACCTTGTATC	3350
3178	Qy	CCAGAATGGGAAGCAGAAAGTGTCCAAGCAGTTTCGGTCTGTCCGGGGCGTGCTATATC	3237
3351	Db	CCAGAATGGGAAGCAGAAAGTGTCCAAGCAGTTTCGGTCTGTCCGGGGGTGGCTATATC	3410
3238	Qy	CTCCGTGTACACGGGTACAAGAGGGGNTATGGAGAGGGTGTGTAAAGTCCCATGAATC	3297
3411	Db	CTTCGGTGTACACGGGTCAAAGAGGGGATATGGAGAGGGCTGCGTAAACGATCCCATGAATC	3470
3298	Qy	GAGAACAATAACAGACGAACTTAAATTTTAAAAAATCTGTGAAGAGAGGAAGTGTATCCAACG	3357
3471	Db	GAGAACATAACAGACGAACTTAAATTTTAAAAAATCTGTGAAGAGAGGAAGTGTATCCACG	3530
3358	Qy	GATACAGGAACGTTGAATGATTTACTTGCACACCAAGGTACAGC-----AGTA	3405
3531	Db	GATACAGGAACGTTGAATGATTTACTTGCACACCAAGGTACAGCTGGATCGCAGATGCA	3590
3406	Qy	TGTAAATCCCGTAAATGCTCGATATAGGATGTCATATGAAGTTGCATCTACAGCATCTGTT	3465
3591	Db	TGTAAATCCCGTAAATGCTCGATATAGGATGTCATATGAAGTTGCATCTACAGCATCTGTT	3650
3466	Qy	AATTACAAACCGACTTATGAAGAAGAAAAGTATACAGATGTACGAAGAGATAATCATTTGT	3525
3651	Db	AATTACAAACCGACTTATGAAGAAGAAAAGTATACAGATGTGAAGAAGAGATAATCATTTGT	3710
3526	Qy	GAATATGACAGAGGGTATGTGAATTTATCCACCTACACGCTGGTTTATATGACAAAGAA	3585
3711	Db	GAATATGACAGAGGGTATGTCAATTTATCCAACCTACACGCTGGTTTATGTCACAAAGAA	3770
3586	Qy	TTAGAATACTTCCAGAAAACCGATAAGGTATGATTTGAGATTCGAGAAAACGGAAGGAAG	3645
3771	Db	TTAGAATACTTCCAGAAAACGATACAGTATGATTTGAGATTCGAGAAAACGGAAGGAAG	3830
3646	Qy	TTTTATTGTAGACAGCGTGGAAATTTACTCCTTTATGGAGGAATAG	3687
3831	Db	TTTTATTGTAGATAGCGTGGAAATTTACTCCTCATGGAAGATAG	3872

PRECINT 6

US-08-100-709-3
Sequence 3, Application US/08100709
Patent No. 5322687
GENERAL INFORMATION:
APPLICANT: Donovan, William P.
APPLICANT: Tan, Yiping
APPLICANT: Jany, Christine S.
APPLICANT: Gonzalez Jr., Jose M.
TITLE OF INVENTION: BACILLUS THURINGIENSIS c1yET4 AND c1yET5
TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
ADDRESSEE: Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,709
FILING DATE: 19930729

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Egolf, Christopher
REGISTRATION NUMBER: 27633
REFERENCE/DOCKET NUMBER: 7205-49
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3934 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 67..3756
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2253..2272
US-08-100-709-3

Query Match 72.98; Score 2687.2; DB 1; Length 3934;
Best Local Similarity 83.54; Pred. No. 0;
Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

Qy 1 TTGACTTCAAAATAGGAAAAATGAGAATGAAATTAATAATGCTTTATCGATTCCACGGTA 60
Db 67 TTGACTTCAAAATAGGAAAAATGAGAATGAAATTAATAATGCTTTATCGATTCCACGGTA 126

Qy 61 TCGAATCATTTCCACACAAATGGATCTATCACAGAGTCTCGTATTGAGGATCTTTTGCT 120
Db 127 TCGAATCTTCCACGCAATGAATCTATCACAGAGTCTCGTATTGAGGATCTTTTGCT 186

Qy 121 ATACCGGAGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCCAACGGGTATT 180
Db 187 GTAGCGGAGGTGAACAATATTGATTCATTTGTTAGCGCATCAACAGTCCCAACGGGTATA 246

Qy 181 AACATAGCTGTAGAACTAGTGTATTAGGCGTACCGTTTGTGTCGACAAATAGCTAGT 240
Db 247 AACATAGCTGTAGAACTATTGGCGGTATTAGTGTGCGCTTTGTGTCGACAACTAGCTAGT 306

Qy 241 TTTTATAGTTTCTTGTTGGTAATTATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC 300
Db 307 TTTTATAGTTTCTTGTTGGGAAATTATGGCCTTAGTGGCAGAGATCCATGGGAAATTTTC 366

Qy 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
Db 367 CTGGAACATGTAGAACAACTTATAAGACAAACAAGTAAACAGAAATACTAGGAATACGGCT 426

Qy 361 CTGCTCGAATTAACAGGTTTAGGAGATTCCTTTTAGAGCTTATCAACAGTCACTTGAAGAT 420
Db 427 ATTGCTCGAATTAGAAAGTCTAGGAAGAGGCTATAGATCTTACCAGCAGGCTCTTTGAAACT 486

Qy 421 TGGCTAGAAACCGTGATGATGCAAGACGAGAGTGTCTTTATACCAATATATAGCC 480
Db 487 TGGTTAGATAACCGAAATGATGCAAGATCAAGAACATTTCTTGAGCGTATGTTGCT 546

Qy 481 TTAGAACTTGATTTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAACAAAGAGTTCCA 540
Db 547 TTAGAACTTGACATTACTACTGCTATACCGCTTTTCAGAAATACGAAATGAAAGAGTTCCA 606

Qy 541 TTATTAATGGTATGCTCAAGCTGCAATTTACACCTATTATTATTGAGAGATGCCCTCT 600
Db 607 TTATTAATGGTATGCTCAAGCTGCAATTTACACCTATTATTATTGAGAGATGCCATCC 666

Qy 601 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTTCAACGTTATTATTGAGCGCAA 660
Db 667 CTTTTTGGTAGTGAATTTGGGGATGCACTCTCCGATGTTAACCAATATTACCAAGAACTA 726

Qy 661 GTGGAAACAAACGAGAGATTTATCCGACTATTTCGGTAGAATGGTATAATACAGTCTTAAT 720
Db 727 ATCAGATATACAGAGGAATTTCTAAACCATTCGTTACAAATGGTATAATACAGGGCTAAAT 786

Qy 721 AGCTTGAGGGGACAAATGCCGAAGTTGGTGGTTTATAATCAATTCCTCGTAGAGATCTA 780
Db 787 AACTTAAGAGGGGACAAATGCTGAAGTTGGTTGGGTATTAATCAATTCCTCGTAGAGACTA 846

Qy 781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCCGACTTATCCA 840
Db 847 ACGTTAGGGGTATTAGATTTAGTAGCCCTATTCCCAAGCTATGATATCTCGCACTTATCCA 906

Qy 841 ATAAATACGAGTGTCTAGTTAAACAAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG 900
Db 907 ATCAATACGAGTGTCTAGTTAAACAAGAGAAATTTATACAGATCCAAATTTGGAGAACAAAT 966

Qy 901 GTAAT-----ATGGCAAGTATGAATTTGGTATAATAATATGACACCTTCGTTTTCGGCT 954
Db 967 GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAATTAATATGACCCATCGTTTTCGTC 1026

Qy 955 ATAGAGACTCGGGTTATCCGAAGCCCGCATCTACTTGATTTCTTAGAACAACTTACAAT 1014
Db 1027 ATAGAGCTGCCATTTTCAGSCCTCCGCATCTACTGATTTTCAGAACAACTTACAAT 1086

Qy 1015 TTTAGCACTTCATCAAGATGGAGTCTACTAGGCAATATGACTTACTGCGGGGCGACACA 1074
Db 1087 TACAGTGCATCAAGCCGTTGGAGTAGCACTCAACATATGAATTTATGGTGGGACATAGG 1146

Qy 1075 ATTCATCTCGGCCCAATAGGAGCGGATTAATAATACCTCAACGCTATGGTCTTACCA---AT 1131
Db 1147 CTTAATCTCCGCCCAATAGGAGGACATTAATACTCAACACAAGGACTTACTTAATAAT 1206

Qy 1132 ACTTCTATTAATCTCTGAAGATTATCATCTTCTCTCGAGACGTATATTGCACTGAATCA 1191
Db 1207 ACTTCAATTAATCTCTGAATTTACAGTTTACGTTTACGTTCTCGAGACGTTTATAGAACAGAAATCA 1266

Qy 1192 ATGCAAGAGTCTCTATGGGGAATTTACCTTGAACCTATTTCATGGTGTCCCTACTGTT 1251
Db 1267 AATGCAAGGACAAATAT-----ACTATTACTCTCTGTGAATGGAGTACCTTGGGCT 1320

Qy 1252 AGATTTAATTTTAGGAACCTCAGAAATACCTTTGAAAGAGGTACTGTCTTAACTATAGTCAA 1311
Db 1321 AGATTTAATTTTATAAACCCCTCAGAAATTTTATGAAGAGGCGCCACTACCTACAGTCAA 1380

Qy 1312 CCTATGAGTACCTCGGGCTTCAATTAAGATTTCAGAACTGAATTAACCAACCAACAAACA 1371
Db 1381 CGGTATCAGGAGTGGGATTCATTTATTTGATTTCAGAACTGAATTAACCAACCAACAAACA 1440

Qy 1372 ACAGAACGACCAAAATTTAGAAATCATATAGTCAATAGTCTTCTACATAGGCTCATTTTCA 1431
Db 1441 ACAGAACGACCAAAATTTAGAAATCATATAGTCAATAGTCTTCTACATAGGACTAATCATA 1500

Qy 1432 CAATCTAGGGTGCATGTACCAAGTATTTCTTGAGCGCACCGTAGTGCAGATCGTACAAT 1491
Db 1501 GGAACACACTTTGAGAGCACCACTTATTTCTTGAGCGCATCGTAGTGCAGATCGTACGAAT 1560

Qy 1492 ACCATTAGTTTCAGATAGCATACACAAATACCATTTGGTAAATCATTTCAACCTTAATTTCA 1551
Db 1561 ACGATTTGGAACCAAAATAGAAATTAACAAATACCATTTGGTAAAGACTGAATCTTCAATTTCA 1620

Qy 1552 GGTACCTCTGTAGTCACTGGCCCGAGGATTTACAGGAGGGGATATAATCCGAATAACGTT 1611
Db 1621 GGTGTACTGTTGTTGAGGGCCAGGATTTACAGGTGGGGATATCTCTGTAGAACAAAT 1680

Qy 1612 AATGGTAGTGTACTAAGTATGGGTCTTAATTTTAATAATACATCATTTACAGCGGTATCCG 1671
Db 1681 ACGGTACATTTTGAGATATACGATTTAAATATAATTTGTCCTATTTCCCAAGATATCCG 1740

Qy 1672 GTGAGAGTGTGTTATGCTCTCTCAACAAATGCTCTCGAGGTAATCTGTCCGAGGGAGT 1731
Db 1741 GTAAGGATCGTTATGCTTCTACTACAGATTTACAATTTTTCACGAGAAATTAATGGAACC 1800

Qy 1732 ACTACTTTTGTATCAAGGATTTCCCTAGTACTATGAGTGAATAATGAGTCTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATTGGTAATTTCTCAAGAACTATCAATAGGGGGGATTAATTTTGAATAATAGA 1860

1792 TCATTAGATTTGCGAATTTCTCTGAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
1861 AGTTTATAGAACTGCGAGATTAGTACTCTCTTTTAAATTTTAAATGCCCCAAGACATTC 1920
1852 ATAGGTATAAGTAATAATCGAGGTAGACAAAGCTTTCTCTTTGATAAATAATGAATTCAT 1911
1921 ACATTGGGGCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1980
1912 CCAATTTACTGCAACCTTTCGAAGCAGAAATACGATTTTGAAGGGCGCGAAGGGCGGTGAAT 1971
1981 CCAGCAGAGGTAACTTTGAGCGAGATATGATTTTGAAGAGCAGCAAAAGGGCGGTGAAT 2040
1972 GCTCTGTTTACTAATCGAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATT 2031
2041 GCTCTGTTTACTTACAAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATT 2100
2032 GATCAAGTATCCAAATTTAGTGGCGGTGTTTATCGGATGAATTTCTGTTAGATGAAAGAGA 2091
2101 GACCAAGTGTCCAAATATGTTGGCATGTTTATCAGATGAATTTTGTGATGAGAAGCGA 2160
2092 GAATTTACTTGAAGAGTGAATATGCGAAACGACTCAGTGTATGAAGAACTTACTCCAA 2151
2161 GAATTTTGAAGAGTGAATATGCGAAGCGACTCAGTGTATGAAGAACTTACTCCAA 2220
2152 GATCCAACTTCACATCCATCAATTAAGCAACAGACTTTCATATCTACTAATGAGCAATCG 2211
2221 GATCCAACTTCACATCCATCAATTAAGTGGCAATTAAGTTTCGCATCCATCGATGGAATCA 2280
2212 AATTTACATCTATCCATGAACTCTGAACATGATGATGTTGGGAGTGAAGAGCAATTACA 2271
2281 AACTTTCCCTCTATTAAATGAGCTATCTGAACATGATGATGTTGGGAGTGAAGAGTGTACC 2340
2272 ATCCAGGAAGGAATGACGTATTTAAAGAGAAATTAAGTCACTACCGGAGCTTTTAAAT 2331
2341 ATTGAGGAAGGAATGACGTATTTAAAGAGAAATTAAGTCACTACCGGAGTGTATTAAAT 2400
2332 GAGTGTATCCGACGTATTATATCAAAATATAGGAGAGTCCGAATTTAAAGCTTTATAT 2391
2401 GAGTGTATCCAAATTTATTTATCAAAATATAGGAGAGTCCGAATTTAAAGCTTTATAG 2460
2392 CCGTACCAATTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGT 2451
2461 CCGTATCAATTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGT 2520
2452 TATAATGCGAAACATGAACTATGATGTTCCAGGTACCGAGTCCGATGCGGTGCTTCA 2511
2521 TACATGCAAGCATGAACATGATGTTCCAGGTACCGATCCGATTCCTATGGCGGCTTCA 2580
2512 GTTGAAGCCCAATCGGAAGTTCGGAAGACCGAATCGATGCGCACCATTTTGAATGG 2571
2581 GTTGAAGCCCAATCGGAAGTTCGGAAGACCGAATCGATGCGCACCATTTTGAATGG 2640
2572 AATCTGATCTAGATTGTTCTGCGAGATGAGAGAAATGTCGCGATCATTTCCCATCAT 2631
2641 AATCTGATCTAGATTGTTCTGCGAGATGAGAGAAATGTCGCGATCATTTCCCATCAT 2700
2632 TTCTCTTTGGATATTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2691
2701 TTCTCTTTGGATATTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2760
2692 GTATTCAAGATTAAAGCAGGAAGGTTCATGCAAGACTAGGGAATCTGGAATTTTATGAA 2751
2761 GTATTCAAGATTAAAGCAGGAAGGTTCATGCAAGATTAGGAATCTGGAATTTATCGAA 2820
2752 GAGAAACCAATTATAGGAGAGCACTGCTCTGTTGAAGAGAGCAGAGAGAGAGAGAGAGAG 2811
2821 GAGAAACCAATTATTTGGAGAGCACTGCTCTGTTGAAGAGAGCAGAGAGAGAGAGAGAGAG 2880
2812 GACAAACGTTGAAACTCAATTTGGAACAAACAGGTATATACAGAGCGCAAAAGAGCT 2871
2881 GACAAACGTTGAAACTCAATTTGGAACAAACAGGTATATACAGAGCGCAAAAGAGCT 2940
2872 GTGGATGCTTTTATTTGTAGATTCTCAATATATAGATTACAGCGGATACAAACATTTGCG 2931

2941 GTGGATGCTTTTATTTCTGATAGTTCTCAATATGATCAATTTACAGCGGATACAAACATTTG 3000
2932 ATGATTCATCGGCGAGATAAATCTGTTTCATCGAATTCGAGAGCTTATCTGTCAGAAATTA 2991
3001 ATGATTCATCGGCGAGATAAATCTGTTTCATCGAATTCGAGAGCTTATCTTTTCAGAAATTA 3060
2992 TCTGTTATCCCGGGTGAATCGCGAAATTTTGAAGAAATTTGAAGTTCGCAATTCATCT 3051
3061 CTTGTTATCCCAAGGTGTAATGCGGAATTTTGAAGAAATTTGAAGTTCGCAATTCATCT 3120
3052 GCAATCTCCCTATACGATCGGAGAAATGTCGTTAAAAATGTTGATTTTAAATATGATTA 3111
3121 GCAATGTCCTTATACGATCGGAGAAATGTCGTTAAAAATGTTGATTTTAAATATGATTA 3180
3112 GCATGCTGGAATGTAAGAGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 3171
3181 ACATGTTGGAATGTAAGAGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3172 GTTATCCCGAGAAATGGAAGCAGAGTGTCAAGACAGTTCGCGTCTGTCGCGGGCGTGGC 3231
3241 GTTATCCCGAGAAATGGAAGCAGAGTGTCAAGACAGTTCGCGTCTGTCGCGGGCGTGGC 3300
3232 TATATCTCTCTGTCACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATTCAT 3291
3301 TATATCTCTCTGTCACAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATTCAT 3360
3292 GAATCGAGAAACATACAGAGCAACTAAAAATTTAAAACTGTGAAGAGAGAGAGAGTAT 3351
3361 GAATCGAGAAACATACAGAGCAACTAAAAATTTAAAACTGTGAAGAGAGAGAGAGTAT 3420
3352 CCAACGGATACAGAAACGTGTAATGATTTACTGACACCAAGGTACAGCAGTATGTAAT 3411
3421 CCAACGGATACAGAAACGTGTAATGATTTACTGACACCAAGGTACAGCAGTATGTAAT 3480
3412 TCCCGTAAATGCTGGATATGAGGATGCATATGAAGTTGATGATGATGATGATGATGATGAT 3471
3481 TCCCGTAAATGCTGGATATGAGGATGCATATGAAGTTGATGATGATGATGATGATGATGAT 3540
3472 AAACCGACTTATGAAGAGAAACGTATACAGATGATGATGATGATGATGATGATGATGATGAT 3531
3541 AAACCGACTTATGAAGAGAAACGTATACAGATGATGATGATGATGATGATGATGATGATGAT 3600
3532 GACAGAGGTATGTAATTTATCCACACTACCGTGGTGTATGATGATGATGATGATGATGATGAT 3591
3601 GACAGAGGTATGTAATTTATCCACACTACCGTGGTGTATGATGATGATGATGATGATGATGAT 3660
3592 TACTTCCAGAAACCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3651
3661 TACTTCCAGAAACCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
3652 GTAGACAGCTGGAAATTTACTCTTTATGAGGAATAG 3687
3721 GTAGATAGCTGGAACTACTCTCTCATGGAAGATAG 3756

RESULT 7

US-08-176-865-3

; Sequence 3, Application US/08176865

; Patent No. 5616319

; GENERAL INFORMATION:

; APPLICANT: Donovan, William P.

; APPLICANT: Tan, Yiping

; APPLICANT: Jany, Christine S.

; APPLICANT: Gonzalez Jr., Jose M.

; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYE4 AND CRYE5

; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.

; ADDRESSEE: Nadel

; STREET: 1601 Market Street, 36th Floor

; CITY: Philadelphia

Db 1621 GGTGTTACTGTTGTTGGAGGGCCAGATTTACAGGTGGGATATCTTCTGTAGACAAAT 1680
Qy 1612 AATGCTAGTACTAGTATGGTCTTAATTTTAATAATACATCATTAACAGCGGTATCGC 1671
Db 1681 ACGGCTACATTTGGAGATATACGATTAATATATGTGCCATTTATCCAAAGATATCGC 1740
Qy 1672 GTGAGATTGCTGTTCTCTCAACAATGCTCTGAGGGTAACTGTCCGAGGGAGT 1731
Db 1741 GTAAGATTCTGTTATGCTTCTACTACAGATTTACAATTTTTCACGAGAAATTAATGGAACC 1800
Qy 1732 ACTACTTTTATCAAGGATTCCTAGTACTATGATGCAATATGAGTCTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATTTGTTAAATTTCTCAAGAACTATGAATGCGGGGATTAATTTTGAATATAGA 1860
Qy 1792 TCAATTTAGATTGCAAGATTTCTCTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
Db 1861 AGTTTATGACCTGAGGATTTAGTACTCTTTTAATTTTAAATGCCCAGCAATTC 1920
Qy 1852 ATAAGTATAAGTAAATATGCAAGGTAGACAAACGTTTCTCTTTGATAAATTTGAATTCAT 1911
Db 1921 ACATTTGGGTGCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGAGTCCGAATTTGTT 1980
Qy 1912 CCNATTACTGCAACTTCGAGCAGAAATACGATTTTAGAAGGGCCAGAGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAAACATTTGAGCAGAAATATGATTTTAGAAGAGCACAAAAGGCGGTGAAT 2040
Qy 1972 GCTCTGTTTACTTAATACGAATCCAAAGAGATTGAACAAGATGTGACAGATTTATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTGAACAAGATGTGACAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGGTGTTTATCGGATGAATTTCTGCTTAGATGAAGAAGA 2091
Db 2101 GACCAAGTGTCCAATATGCTGGCATGTTTATCAGATGAATTTTCTTGATGAGAAGCGA 2160
Qy 2092 GAATTTACTTGGAAAGTGAATATCGGAAACGACTCAGTGTAGTGAAGAACTTACTCAA 2151
Db 2161 GAATTTATTTGAGAAGTGAATATGCGAAGCGACTCAGTGTAGTGAAGAACTTACTCAA 2220
Qy 2152 GATCCAACTTCCATCATCAATGAAGCAACAGACTTTCATATCTACTTAATGAGCAATCG 2211
Db 2221 GATCCAACTTCCATCATCAGTGGGCAATTAAGTTTCGATCATCGATGGACAATCA 2280
Qy 2212 AATTTACATCTATCCATGAACAACTCTGAACATGGATGGTGGGAAGTGAAGAACATTTACA 2271
Db 2281 AACTTCCCTCTATTAATGAGCTATCTGAACATGATGGTGGGAAGTGGCAATGTTTACC 2340
Qy 2272 ATCCAGGAAGGAATGAGTATTTAAGAGAAATTTACGTACACTACCGGGGACTTTTAAT 2331
Db 2341 ATTCAAGGAAGGAATGAGCTATTTAAGAGAAATTTACGTACACTACCGGGGACTTTTAAT 2400
Qy 2332 GAGTGTATCCGAGCTATTTATATCAAAAAATAGGAGTGGGAATTAAGAGCTTTATACT 2391
Db 2401 GAGTGTATCCAAATTTATATCAAAAAATAGGAGTGGGAATTAAGAGCTTTATAGC 2460
Qy 2392 CGCTACCAATTTAAGAGGTATATTTGAAGATAGTCAAGATTTTAGAGATATATTTGATTCGT 2451
Db 2461 CGCTATCAATTTAAGAGGTATATTTGAAGATAGTCAAGATTTTAGAGATATATTTGATTCGT 2520
Qy 2452 TATTAATGCAACATGAACATTTGATGTTCCAGTACCGAGTCCGATGATGCGCGTTTCA 2511
Db 2521 TACAATGCAACATGAACATTTGATGTTCCAGTACCGAGTCCGATGATGCGCGTTTCA 2580
Qy 2512 GTTGAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGACCACTTTTGAATGG 2571
Db 2581 GTTGAAGCCCAATCGGAAGTGGGAGAACCAATCGATCGACCACTTTTGAATGG 2640
Qy 2572 AATCCTGATCTAGATTTGTTCTCGAGAGATGGAGAAATGTCGCAATTCATCCCATCAT 2631
Db 2641 AATCCTGATCTAGATTTGTTCTCGAGAGATGGAGAAATGTCGCAATTCATCCCATCAT 2700
Qy 2632 TTCTCTTTGGATATGATTTGATGTCACAGACTTGCATGAGATCTTAGGCGTGTGGGTG 2691
Db 2701 TTCATTTGGATATTTGATTTGGGTGTCACAGACTTGCATGAGAACCTTAGGCGTGTGGGTG 2760

Qy 2692 GTATTTCAAGATTTAAGACGACGAGAAAGTTCATGCAAGACTAGGSAATCTCGAAATTTTATCGAA 2751
Db 2761 GTATTTCAAGATTTAAGACGACGAGAAAGTTCATGCAAGACTAGGSAATCTCGAAATTTTATCGAA 2820
Qy 2752 GAGAAACCAATTTATAGGAGAAAGCACTGTCTCTGTGTGAAGAGAGACGAGAAAGAAATTTGAGAA 2811
Db 2821 GAGAAACCAATTTATTTGAGAGAGCACTGTCTCTGTGTGAAGAGAGCGGAAAGAAATTTGAGAA 2880
Qy 2812 GACAAACGTGAAACCTACAAATTTGGAACAAACAAACGAGTATATACAGAGCGCAAAAGAGCT 2871
Db 2881 GACAAACGGGAAACCTACAAATTTGGAACAAACAAACGAGTATATACAGAGCGCAAAAGAGCT 2940
Qy 2872 GTGATGCTTTATTTGTTAGATTTCTCAATATATATAGATTTTACAGCGGATACAAACATTTGCGC 2931
Db 2941 GTGATGCTTTATTTCTGATGATTTCTCAATATGATCAATTTTACAGCGGATACAAACATTTGCGC 3000
Qy 2932 ATGATTCATGCGGAGATATAAATTTGTTTCATCGAAATTCGAGAGGCTTATCTGTGAGAAATTA 2991
Db 3001 ATGATTCATGCGGAGATATAAATTTGTTTCATCGAAATTCGAGAGGCTTATCTTTTCAGAAATTA 3060
Qy 2992 TCTGTTATCCCGGTTGTAATGCGGAAATTTTTCGAGAAATTTTTCGAGGTCGCATTTACT 3051
Db 3061 CCTGTTATCCAGGTGTAAATGCGGAAATTTTTCGAGAAATTTTTCGAGGTCGCATTTACT 3120
Qy 3052 GCAATCTCCCTATACGATGCGAGAAATGTCGTTAAAAATGCTGATTTTAAATTAATGGAATTA 3111
Db 3121 GCAATGCTCTTATACGATGCGAGAAATGTCGTTAAAAATGCTGATTTTAAATTAATGGAATTA 3180
Qy 3112 GCATGCTGGAATTTAAAGGGCATGTAGATGTAACAAGAGCCATCAACGCTTCTGTCTCT 3171
Db 3181 ACATGTTGGAATTTAAAGGGCATGTAGATGTAACAAGAGCCATCAACGCTTCTGTGACCTT 3240
Qy 3172 GTTATCCAGNAATGGGAAGCAAGATGTCAACAAGCTTCCGCTCTGTCCGGGGCGTGGC 3231
Db 3241 GTTATCCAGNAATGGGAAGCAAGATGTCAACAAGCTTCCGCTCTGTCCGGGGCGTGGC 3300
Qy 3232 TATATCTCTCGTGTCAACAAGCGATACAAAGAGGATATGGAGAGGTTGTTAAACGATCCAT 3291
Db 3301 TATATCTCTCGTGTCAACAAGCGATACAAAGAGGATATGGAGAGGCTGCGTAAACGATCCAT 3360
Qy 3292 GAAATCGAGAACAAATACAGACGAACTAAAAATTTAAAACTGTGGAAGAGAGAAAGTGTAT 3351
Db 3361 GAAATCGAGAACAAATACAGACGAACTAAAAATTTAAAACTGTGGAAGAGAGAAAGTGTAT 3420
Qy 3352 CCAACGATACAGAAACGTGTATGATTTACTCCACACCAAGGTACAGCAGTATGTAAT 3411
Db 3421 CCAACGATACAGAAACGTGTATGATTTACTCCACACCAAGGTACAGCAGTATGTAAT 3480
Qy 3412 TCCCGTAAATGCTGATATGAGGATGATGAAAGTTCATCTACAGCATCTGTTAAATTAAC 3471
Db 3481 TCCCGTAAATGCTGATATGAGGATGATGAAAGTTCATCTACAGCATCTGTTAAATTAAC 3540
Qy 3472 AAACCGATTTATGAAGAGAAACGTTATACAGATGTACGAAGATAATCAATTTGGAATTAAT 3531
Db 3541 AAACCGATTTATGAAGAGAAACGTTATACAGATGTACGAAGATAATCAATTTGGAATTAAT 3600
Qy 3532 GACAGAGGATGTGAATTTATCCACCACTACCAGTCCGTTATATGACAAAGAAATTAGAA 3591
Db 3601 GACAGAGGATGTGAATTTATCCACCACTACCAGTCCGTTATATGTAACAAAGAAATTAGAA 3660
Qy 3592 TACTTCCAGAAACCGAATTAAGGATGATGAAATTTGAGATTTGGAGAAACGGAAGGAAATTTAT 3651
Db 3661 TACTTCCAGAAACCGAATTAAGGATGATGAAATTTGAGATTTGGAGAAACGGAAGGAAATTTAT 3720
Qy 3652 GTAGACAGCGTGAATTTACTCTCTTATCGAGGAATAG 3687
Db 3721 GTAGATAGCGTGAATTTACTCTCTCTCATGGAAGAAATAG 3756

RESULT 8
US-08-474-038-3
; Sequence 3, Application US/08474038

Query Match	Best Local Similarity	Score	2687.2;	DB 1;	Length	3934;
Matches	3088;	Conservative	0;	Mismatches	593;	Indels
						Gaps
1	TTGACCTTCAATAGGAAAAATGAGAAATGAAATATATAAATGCTTTTATCGATTCCAGCTGTA	60				
67	TTGACCTTCAATAGGAAAAATGAGAAATGAAATATATAAATGCTTTTATCGATTCCAGCTGTA	126				
61	TCGAATCATTCACAAATGGATCTATCACCAGATGCTCGTATTGAGAGATCTTTGGT	120				
127	TCGAATCCTTCCACGAAATGAAATCTATCACCAGATGCTCGTATTGAAAGATGCTTTGGT	186				
121	ATAGCCGAGGGGAATATATCAATCCACTTTGTTAGGCGATCAACAGTCCAAACGGGTATT	180				
187	GTAGCCGAGGTGAACATATTGATCCATTTGTTAGGCGATCAACAGTCCAAACGGGTATA	246				
181	AACATAGCTGGTGAATATCTAGTGTGTTATAGGCGTACCGTTTGTGCGAACAAATAGCTAGT	240				
247	AACATAGCTGGTGAATATTTGGGCGTATTAGGTGTGCGGTTTGTGCGAACAACTAGCTAGT	306				
241	TTTTATAGTTTTCTTCTGTTGGGAATTTATGSCCTAGTGGCAGAGATCCATGGGAAATTTTC	366				

Db 1381 CCGTATCAGGAGTTGGGATTCATATTTTGAATTCAGAAATCGAATATCCACAGAAACA 1440
Qy 1372 ACAGAACGCCAAATTTATGAATCATATAGTTCATAGGTTATCTCACATAGGGCTCATTTTCA 1431
Db 1441 ACAGAACGCCAAATTTATGAATCATATAGTTCATAGGTTATCTCATATAGGACTATCATATA 1500
Qy 1432 CAATCTAGGGTGCATGTACAGTATATCTTGGAGCCACCGTAGTGCAGATCGTCAAAAT 1491
Db 1501 GGAACACATTTGAGAGCACCACTCTATCTTGGAGCCACCTAGTGCAGATCGTACGAAT 1560
Qy 1492 ACCATTAGTTCAGATAGCATAACACAAATACCAATTCGTTAAATCATTCACACCTTAAATCA 1551
Db 1561 AGATTGGACCAATAGAAATTAACAATAACCATTTGGTAAAGCACTGAATCTTCATTTCA 1620
Qy 1552 GGTACCTCTGTAGTCCAGTCCAGGATTTACAGGAGGGGATATTAATCCGAACTAACGTT 1611
Db 1621 GGTGTTACTGTTTGGAGGGCCAGGATTTACAGGTGGGGATCTCTTCGTAGAACAAAT 1680
Qy 1612 AATGGTAGTGTACTAAGTAGTGGTCTTAATTTTAAATAATACATTAATACAGCGGTATCGC 1671
Db 1681 ACGGGTACATTTGGAGATATACGATTAATAATTAATGTGCCATTAATCCCAAGATATCGC 1740
Qy 1672 GTGAGAGTTCGTTATGCTCTCTCAACAAATGGTCTCAGGGTAACGTGCGAGGGAGT 1731
Db 1741 GTAAGATTCGTTATGCTCTCTACTACAGATTTACAATTTTTCACAGAAATTAATGGAACC 1800
Qy 1732 ACTACTTTTGTCAAGGATTCCTCTAGTACTATGAGTGCAAATGAGTCTTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATTTGGTAAATTTCTCAAGACTATGAATAGGGGGATTAATTTAGAAATATAGA 1860
Qy 1792 TCATTTAGATTTGCGAATTTCTGTAGGTATTAATGTGCATCTGGCAGTCAAACTGCTGGA 1851
Db 1861 AGTTTTAGAACTGCGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCCAAGCACATTC 1920
Qy 1852 ATAAGTATAAGTAAATAATCCAGGTAGACAAAGCTTTTCACTTTGTATAAAATTTGAATTCATT 1911
Db 1921 ACATTTGGGTGCTCAGAGTTTTTCAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1980
Qy 1912 CCAATTAATCTGCAACCTTGAAGCAGAAATACAGATTTAGAAAGGGCGCAAGAGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAACTTTGAGCGCAGATATGATTTAGAAAGCAGCAAAAGCGGTGAAT 2040
Qy 1972 GCTCTGTTTATTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATTCGCTTAGATGAAGAAAGAGA 2091
Db 2101 GACCAAGTGTCCAAATATGTTGGCATGTTTTATCAGATGAATTTTGTGATGAGAAAGCGA 2160
Qy 2092 GAATTAATCTGAGAAAGTGAATATATCGAAACGACTCAGTGTAGAAAGAAACTTACTCCAA 2151
Db 2161 GAATTAATTTGAGAAAGTGAATATATCGAAGCGACTCAGTGTAGAAAGAAACTTACTCCAA 2220
Qy 2152 GATCCAAACTTCAATCCATCAATTAAGCAACAGACTTTCATATCTACTAATAGCAATCG 2211
Db 2221 GATCCAAACTTCAATTCATCAGTGGCAATTAAGTTTCGCATCCATCGATGCAATCA 2280
Qy 2212 AATTTCAATCTATCCATGAACAACTGTGAACATGATGTTGGGAGGTGAGAACATTAACA 2271
Db 2281 AACTTTCCCTCTTAATTAATGAGCTATCTGAACATGATGTTGGGAGGTGCGAATTTTACC 2340
Qy 2272 ATCCAGGAGGAAATGACCTATTTAAAGAGAAATTAAGTACACTACCGGGGACTTTTAAT 2331
Db 2341 ATTGAGGAAGGAATGACCTATTTAAAGGAATTAAGTACACTACCGGGGACTTTTAAT 2400
Qy 2332 GAGTGTATCCGACGTATTTATATCAAAAAATPAGGAGGTGCGAATTTAAAGCTTTATACT 2391
Db 2401 GAGTGTATCCAAATTTATATCAAAAAATPAGGAGGTGCGAATTTAAAGCTTTATAGC 2460
Qy 2392 CGCTACCAATTAAGAGGTATATTGAAGATGTCAAGATTTAGAGATATATTGATTCGT 2451
Db 2461 CGCTATCAATTAAGAGGTATATTGAAGATGTCAAGATCTAGAGATTTATTTAATTCGT 2520

Qy 2452 TATAATCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGTATGGCGCTTTTCA 2511
Db 2521 TACAATGCAAGCATGAACATTTGATGTTCCAGGTACCGATTCCTATGGCGCTTTCA 2580
Qy 2512 GTTGAAGCCCAATCGGAAGGTGCGGAAACCGAATCGATCGACCAATTTTGAATGG 2571
Db 2581 GTTGAAGCCCAATCGGAAGGTGCGGAAACCGAATCGATCGACCAATTTTGAATGG 2640
Qy 2572 AATCTGATCTAGATTTGTTCTTCGACAGATGAGAAAATGTCGCCATCATTTCCCATCAT 2631
Db 2641 AATCTGATCTAGATTTGTTCTTCGACAGATGAGAAAATGTCGCCATCATTTCCCATCAT 2700
Qy 2632 TTCTCTTTGGATATTTGATTTGGATGTCACAGATTTGCATGAGAACTTAGGCCGTGGGGTG 2691
Db 2701 TTCACTTTGGATATTTGATTTGGGTGTCACAGATTTGCATGAGAACTTAGGCCGTGGGGTG 2760
Qy 2692 GTATTTCAAGATTTAAGACCCAGAGAGGTCTATGCAAGACTAGGAAATCTGGAATTTATGAA 2751
Db 2761 GTATTTCAAGATTTAAGACCCAGAGAGGTCTATGCAAGATTTAGGAAATCTGGAATTTATGAA 2820
Qy 2752 GAGAAACCATTTATAGGAGAGCACTGCTCTGTGTAAGAGAGCAGAGAAAATTTGGAGA 2811
Db 2821 GAGAAACCATTTATTTGGAGAGCACTGCTCTGTGTAAGAGAGCAGAGAAAATTTGGAGA 2880
Qy 2812 GACAAACGTGAAACACTACAATTTGAAACAAAAACGAGTATATACAGAGCAAAAGAGCT 2871
Db 2881 GACAAACGTGAAACACTACAATTTGAAACAAAAACGAGTATATACAGAGCAAAAGAGCT 2940
Qy 2872 GTGATGCTTTATTTGTAGATTTCTCAATATATATAGATTTCAAGCGGATACAAACATTTGGC 2931
Db 2941 GTGGATGCTTTATTTGCTAGATTTCTCAATATATATCAATTTACAAGCGGATACAAACATTTGGC 3000
Qy 2932 ATGATTCATGCGGAGATTAATCTGTTCAATCGAATTTGAGAGGCTTATCTGTCAGAAATTA 2991
Db 3001 ATGATTCATGCGGAGATTAATCTGTTCAATCGAATTTGAGAGGCTTATCTGTCAGAAATTA 3060
Qy 2992 TCTGTTATTCGGGTGTTAAATCGGAAATTTTGAAGAAATTTAGAAAGTTCGCAATTTACT 3051
Db 3061 CTGTTATTCAGAGGTGTTAAATCGGAAATTTTGAAGAAATTTAGAAAGTTCGCAATTTACT 3120
Qy 3052 GCAATCTCCCTATACGATCGGAGAAATGCTGTTAAAAATGTTGATTTTAAATTAATGGAATTA 3111
Db 3121 GCAATGCTCTTATACGATCGGAGAAATGCTGTTAAAAATGTTGATTTTAAATTAATGGAATTA 3180
Qy 3112 GCATGCTGGAATGTAAGAGGCATGTAGATCTCAACAGAGCCATCAAGTTCTGTCCTT 3171
Db 3181 ACATGTTGGAATGTAAGAGGCATGTAGATCTCAACAGAGCCATCAAGTTCTGTCCTT 3240
Qy 3172 GTTATCCAGAAATGGGAAGCAGAAAGTGTCAACAGAGTTCGCGTCTGTCGGGGGCTGGC 3231
Db 3241 GTTATCCAGAAATGGGAAGCAGAAAGTGTCAACAGAGTTCGCGTCTGTCGGGGGCTGGC 3300
Qy 3232 TATATCTCTCGTGTACAGCGTACAAAGAGGATATGAGAGGGTGTGTAACGATCCAT 3291
Db 3301 TATATCTCTCGTGTACAGCGTACAAAGAGGATATGAGAGGGTGTGTAACGATCCAT 3360
Qy 3292 GAAATCGAGAACATACAGACCAACTAAAAATTTAAAACTGTGAAGAGAGGAAGTGTAT 3351
Db 3361 GAAATCGAGAACATACAGACCAACTAAAAATTTAAAACTGTGAAGAGAGGAAGTGTAT 3420
Qy 3352 CCAACGGATACAGAAACGTGTAATGATTTACTGCACCAAGGTGTACAGAGTATGTAAT 3411
Db 3421 CCAACGGATACAGAAACGTGTAATGATTTACTGCACCAAGGTGTACAGAGTATGTAAT 3480
Qy 3412 TCCCGTAATGCTGGATATGAGGATGCAATATGAAGTTGATCTACAGCATCTGTTAATTTAC 3471
Db 3481 TCCCGTAATGCTGGATATGAGGATGCAATATGAAGTTGATCTACAGCATCTGTTAATTTAC 3540
Qy 3472 AAAACGGATTTAGAGAGAAACGTATACAGATGTACAGAGATTAATCATTTGGAATAT 3531
Db 3541 AAAACGGATTTAGAGAGAAACGTATACAGATGTACAGAGATTAATCATTTGGAATAT 3600

QY 3532 GACAGAGGGTATGTGAATTATCACCACCTACACAGCTGGTTATATGACAAAAGAAATTAGAA 3591
DB 3601 GACAGAGGGTATGTGAATTATCACCAGTACACAGCTGGTTATGTGACAAAAGAAATTAGAA 3660
QY 3592 TACTTCCAGAAACCGATAGGATGATGGATTGGAGAAACCGAAGGGAAGTTATT 3651
DB 3661 TACTTCCAGAAACAGATAGGATGATGGATTGGAGAAACCGAAGGGAAGTTATT 3720
QY 3652 GTAGACAGCTGGAAATCTACTCTTATGGAGGAATAG 3687
DB 3721 GTAGATAGCTGGAAATCTACTCTTATGGAGGAATAG 3756

RESULT 9
US-08-779-046-3
; Sequence 3, Application US/08779046
; Patent No. 5854053
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yiping
; APPLICANT: Jan, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESS: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,046
; FILING DATE: 06-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3934 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 67..3756
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2253..2272
US-08-779-046-3

Query Match 72.9%; Score 2687.2; DB 2; Length 3934;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;
QY 1 TTGACTTCAATAGGAAATGAGATTAATTAATGCTTTATCGATTCCAGCTGTA 60
DB 67 TTGACTTCAATAGGAAATGAGATTAATTAATGCTTTATCGATTCCACGGTA 126

QY 61 TCGAATCAITTCACACAAATGGATCTATCACCAGATGCTCGTATTGAGGATCTTTGTGT 120
DB 127 TCGAATCTTTCCACGCAAAATGAATCTATCACCAGATGCTCGTATTGAGGATCTTTGTGT 186
QY 121 ATAGCCAGAGGGGAATATATCAATCCACTTTGTAGCGCATCAACAGTCCAAACGGGTATT 180
DB 187 GTAGCCAGGTGAACAATATTGATCCATTTGTTAGCGCATCAACAGTCCAAACGGGTATA 246
QY 181 AACATAGCTGTAGATATCTAGGTGATTATAGGCGGTACCGTTTCTGCTGGACAAATTTTC 240
DB 247 AACATAGCTGTAGATATTTGGGCGGTATTAGGTGTGCCGTTTCTGCTGGACAACTAGCTAGT 306
QY 241 TTTTATAGTTTTCTTTGTTGGTGAATTTATGGCCCGCGGCGAGAGATCAGTGGGAAATTTTC 300
DB 307 TTTTATAGTTTTCTTTGTTGGGGAATTTATGGCTAGTGGCAGAGATCCATGGGAAATTTTC 366
QY 301 CTAGAACATGTGCAACAACTTTATTAATCAACAAATACAGAAATGCTAGGAAATACGGCA 360
DB 367 CTGGAAACATGTAGAACAACTTTATAGAACAACTAGTAAACAGAAATATCTAGGAATACGGCT 426
QY 361 CTGCTCGATTACAGGTTTATAGGAGATTCTTTTAGAGCTATCAACAGTCACTTTGAAGAT 420
DB 427 ATTGCTCGATTAGAGGTTCTAGGAAGGCTATAGATCTTTACCAGAGGCTCTTTGAAACT 486
QY 421 TGGCTAGAAAAACCGTGTATGATGCAAGAACGAGAAAGTGTCTTTTATACCAATATATAGCC 480
DB 487 TGGTTAGATAACCGGAATGATGCAAGATCAAGAAAGCATTTATCTTGAGCGCTATGTTGCT 546
QY 481 TTAGAACTTGATTTCTTAATGCCAGTCCGCTTTTCGCAATTTAGAAACCAAGAAATTTCCA 540
DB 547 TTAGAACTTGATTTCTTAATGCCAGTCCGCTTTTCGCAATTTAGAAACCAAGAAATTTCCA 606
QY 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCCTCT 600
DB 607 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCCTCT 666
QY 601 CTTTTTGGTGTGAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATTGAGCGCCAA 660
DB 667 CTTTTTGGTGTGAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATTGAGCGCCAA 726
QY 661 GTGGAACAAACGAGAGATTATTCCGACTATTCGCTAGAAATGGTATAATACAGGCTTAAT 720
DB 727 ATCAGATATACAGAGGATTTCTAACCATTCGCTACAAATGGTATAATACAGGCTTAAT 786
QY 721 AGCTTGAGAGGGACAAATCCGCAAGTGGGTGCGTTATTAATCAATTCGCTAGAGATCTA 780
DB 787 AACTTAAGAGGGACAAATGCTGAAAGTTGGTTGCGGTATAATCAATTCGCTAGAGACCTA 846
QY 781 ACCTTAGGGGTATTAGATCTAGTGGCAGCTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
DB 847 ACCTTAGGGGTATTAGATTTTAGTCCCTATTTCCCAAGCTATGATGATCTCGCACTTATCCA 906
QY 841 ATAAATACGAGTCTCAGTTTAAACAGGGGAAGTTTATACAGAGCAATTTGAGGCAACAGGG 900
DB 907 ATCAATACGAGTCTCAGTTTAAACAGGAAATTTATACAGATCCAAATTTGGGAGAACAAAT 966
QY 901 GTAAAT-----ATGGCAAGTGAATTTGGTATAATAATTAATGACCTTCGTTTTCGGCT 954
DB 967 GCACCTTCAGGATTTGCAAGTACGAAATTTGTTTAAATAATGACCAATCGTTTTCGCT 1026
QY 955 ATAGAGACTCGGTTATCCGAAAGCCGCACTCTACTTTGATTTTCTAGAACAACTTACAAT 1014
DB 1027 ATAGAGGCTGCAATTTTCAGGCTCCGCACTCTACTTTGATTTTCCAGAACAACTTACAAT 1086
QY 1015 TTTTAGCACTTCATCAGATGGAGTGTACTAGGATATGACTTACTTGGCGGGGGCACACA 1074
DB 1087 TACAGTGCATCAAGCCCGTTGGAGTAGCACTCAACATATGAATTAATTTGGGTGGGACATAGG 1146
QY 1075 ATTCAATCTCGGCAATAGGAGGGGGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1131
DB 1147 CTAACTTCCGCCCAATAGGAGGACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1206

1132 ACTCTTAATTAATCCGTGAAGATTATCAATTTCTCTCGAGACGTATATATTGGACTGAATCA 1191
1207 ACTTCAAATTAATCCGTGAACATTAACAGTTTACGTCTCGAGACGTATATAGAACAGAAATCA 1266
1192 TATGACGAGTCTCTTATAGGGGAATTTACCTTTGAACCTATTTATCATGGTGTCCCTACTGTT 1251
1267 AATGACGGAGCAAAATAT-----ACTATTTACTCTCTGTGTAATGGAGTAGTACCTTTGGGCT 1320
1252 AGATTTAATTTTATGAAACCTCTCAGAAATACCTTTTGAAGAGGTACTGTCTAACTATATAGTCAA 1311
1321 AGATTTAATTTTATAAACCTCTCAGAAATATTTATGAAGAGGGCCCACTACTACAGTCAA 1380
1312 CCCTATAGTACCTCTGGGCTTCAATTAAGAAATTCAGAAATTCAGAAATTCACCAAGAGACA 1371
1381 CCGTATCAGGAGTGGGATTCATTAATTTGATTCAGAAATTCAGAAATTCACCAAGAGACA 1440
1372 ACAGAACGACCAAAATATGAATCATATAGTACATAGTTATCTCATATAGGCTCATTTTCA 1431
1441 ACAGAACGACCAAAATATGAATCATATAGTACATAGTTATCTCATATAGGCTCATTTTCA 1500
1432 CAATCTAGGCTGATGTATACAGTATATTTCTTGGACGCCAGGTAGTGCAGATCGTACAAAT 1491
1501 GGAACACCTTTGAGAGCCAGCTCTATTTCTTGGACGCATCGTAGTGCAGATCGTACGAAT 1560
1492 ACCATTAGTTCAGATAGCATAACAAATACCAATTCATTTGGTAAATCATTCACACCTTAATTC 1551
1561 ACGATTGACCAAAATAGAAATTAACAAATACCAATTCATTTGGTAAAGCACTGAATCTTCATTTCA 1620
1552 GGTACCTCTGTAGTCAGTGGGCCAGGATTTACAGGAGGGATATTAATCCGAACTTAACGTT 1611
1621 GGTGTACTGTGTGGAGGGCCAGGATTTACAGGTGGGGATATCTTCGTAGAACAAAT 1680
1612 AATGTAGTGTACTAAGTATGGTCTTAATTTTAAATAATACATCATATACAGCGGTATCGC 1671
1681 ACGGTACATTTGGAGATATACGATTAATATATATATGTGCCATATATCCCAAGATATCGC 1740
1672 GTGAGATTCGTATGCTGCTTCTCAAAATAGTCTCGAGGGTAACTGTGCGAGGGAGT 1731
1741 GTAAGGATTCGTATGCTTCTACTACAGATTTACAAATTTTACAGAAATTAATGGAACC 1800
1732 ACTACTTTTGATCAAGGATTCCTAGTACTATGATGCAAAATGATGCTTTTGACATCTCAA 1791
1801 ACTGTTAATATGGTAAATTTCTCAAGAACTATGAATAGGGGGATTAATTTTGAATATAGA 1860
1792 TCATTTAGATTCAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
1861 AGTTTATAGATTCAGAGATTTAGTACTCTCTTTAATTTTAAATGCCCAGCAATTC 1920
1852 ATAGTATAGTAAATATGACAGGTAGACAAACGTTTCTCTTTGATAAATTTGAATTCATTT 1911
1921 ACATTTGGGTGCTCAGAGTCTTTTCAAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1980
1912 CCAATTTAGTAACTTCGAGAGCAATACGATTTTGAAGGGCCAGAGAGCGGTGAAT 1971
1981 CCAGCAGAGGTAACTTTGAGCAGAAATATGATTTTGAAGAGGACAAAGAGCGGTGAAT 2040
1972 GCTCTGTTTAAATATCAAGATTCAGAAAGATTTGAACAGATGTGACAGATTTATCATATTT 2031
2041 GCTCTGTTTACTTCTACAAATCCAGAGATTTGAACAGATGTGACAGATTTATCATATTT 2100
2032 GATCAAGTATCAATTTATGCGGTGTTTATCGGATGAATTTCTGTTAGATGAAGAGAGA 2091
2101 GACCAAGTGTCCAAATATGCTGCAATTTTATCAGATGAATTTTGTGATGAGAGCGA 2160
2092 GAATTTCTTGAAGAGTGAATATGCGAAACGACTCAGTATGAAGAAACTTACTCCAA 2151
2161 GAATTTATTTGAAGAGTGAATATGCGAAGCGACTCAGTATGAAGAAACTTACTCCAA 2220
2152 GATCCAAACTTCAATTCATCAATGAAGCAACAGACTTTCATATCTACTTAATGAGCAATCG 2211
2221 GATCCAAACTTCAATTCATCAGTGGGCAATTAAGTTTTCGATCCATCGATGGCAATCA 2280
2212 AATTTACATCTATCCATGAACAAATCTGAACATGGATGGTGGGGAAGTGAGAACATTTACA 2271

2281 AACTTCCCTCTATTTAATGAGCTATCTGAAATGATGGTGGGAAGTGCAGATGCTTACC 2340
2272 ATCCAGGAGAAATGAGTATTTTAAAGAGAAATTTACGTCACTACCGGGGACTTTTAAT 2331
2341 ATTCCAGGAGGAAATGACGTATTTTAAAGAGAAATTTACGTCACTACCGGGTACTTTTAAT 2400
2332 GAGTGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCCGAAATTTAAAGCTTATACT 2391
2401 GAGTGTATCCAAATTTATATCAAAAAATAGGAGAGTCCGAAATTTAAAGCTTATAGG 2460
2392 CGCTACCAATTTAAGAGGGTATTTTGAAGATAGTCAAGATTTTGAAGATATATTTGATCGT 2451
2461 CGCTATCAATTTAAGAGGGTATTTTGAAGATAGTCAAGATCTAGAGATTTATTTAATTCGT 2520
2452 TATATCGGAACATGAAACATTCGATGTTCCAGGTACCGAGTCCGATGCGCGCTTCA 2511
2521 TACATGCAAAAGCATGAAACATTTGGATGTTCCAGGTACCGATTTCCCTATGCGCGCTTCA 2580
2512 GTTGAAGCCCAATCGGAAGGTGCGGAAACCGAATCGATCGCACCACTATTTTGAATGG 2571
2581 GTTGAAGCCCAATCGGAAGGTGCGGAAACCGAATCGATCGCACCACTATTTTGAATGG 2640
2572 AATCTGATCTAGATTTGTTCTCTGAGAGATGGAGAAATTTGCGGCATCATTTCCCATCAT 2631
2641 AATCTGATCTAGATTTGTTCTCTGAGAGATGGAGAAATTTGCGGCATCATTTCCCATCAT 2700
2632 TTCTCTTGGATATTTGATTTGATGTCACAGACTTGCATGAGATCTTAGGCGTGTGGGTG 2691
2701 TTCACTTTGGATATTTGATTTGTTGGGTGTCACAGCTTGCATGAGAACCTTAGGCGTGTGGGTG 2760
2692 GTATTTCAAGATTAAGACGCGAGAAAGTTCATCAAGACTTAGGAAATCTCGAAATTTATGAA 2751
2761 GTATTTCAAGATTAAGACGCGAGAAAGTTCATCAAGACTTAGGAAATCTCGAAATTTATGAA 2820
2752 GAGAAACCATTTATTTAGGAGAAAGCTGCTCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2811
2821 GAGAAACCATTTATTTAGGAGAAAGCTGCTCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2880
2812 GACAAACGTTGAAACATCTACAAATTCGAAACAAACGAGTATATACAGAGCAAAAGAGAGCT 2871
2881 GACAAACGTTGAAACATCTACAAATTCGAAACAAACGAGTATATACAGAGCAAAAGAGAGCT 2940
2872 GTGATGCTTTATTTGATGATTTCTCAATATATAGATTTCAAGGCGATACAAACATTTGCG 2931
2941 GTGATGCTTTATTTGATGATTTCTCAATATGATCAATTTACAGGCGGATACAAACATTTGCG 3000
2932 ATGATTCATGCGGAGATTAATCTTTGATTCATCAATTTGAGAGGCTTATCTGTGAGAAATTA 2991
3001 ATGATTCATGCGGAGATTAATCTTTGATTCATCAATTTGAGAGGCTTATCTTTTCAAGAAATTA 3060
2992 TCTGTTATCCCGGTGTAAATGCGGAAATTTTTCGAAAGATTTAGAGGTGCGATTTATCACT 3051
3061 CCTGTTATCCCGGTGTAAATGCGGAAATTTTTCGAAAGATTTAGAGGTGCGATTTATCACT 3120
3052 GCAATCTCCCTATACGATGCGAGAAATGTCGTTTAAATGCTGATTTTAAATGAGATTA 3111
3121 GCAATGCTCTTATACGATGCGAGAAATGTCGTTTAAATGCTGATTTTAAATGAGATTA 3180
3112 GCATGCTGGAATGTAAAGGGCATGTAGATTAACAGAGCCATTCACCGTTCGTCTCTT 3171
3181 ACATGTTGGAATGTAAAGGGCATGTAGATTAACAGAGCCATTCATCGTCTTCGACCTT 3240
3172 GTTATCCAGAGTGGAGAGCAAGATGTCACAAAGCTTCCGCTGTCTCCGGGGCGTGGC 3231
3241 GTTATCCAGAGTGGAGAGCAAGATGTCACAAAGCTTCCGCTGTCTCCGGGGCGTGGC 3300
3232 TATATCTCTCGGTGTCACAGCGTACAAAGAGGATTTAGAGAGGCTTTGTGAACCATTCAT 3291
3301 TATATCTCTCGGTGTCACAGCGTACAAAGAGGATTTAGAGAGGCTTTGTGAACCATTCAT 3360
3292 GAAATCGAGAGCAATACAGAGCAATCTAAAAATTTAAAAATCTGTGAAGAGAGAGAGTGTAT 3351

Db 967 GCACCTTCAGGATTTGCAAGTACGAATTGGTTTAATAATAATGACCAATCGTTTTCTGCC 1026
Qy 955 ATAGAGACTCGCGTTATCCGAAGCCGCACTCTACTTGATTTTCTAGAAACAACCTCAAAAT 1014
Db 1027 ATAGAGGCTGCCATTTTCAGGCTCCGCACTCTACTTGATTTTCCAGAACCACTTACAAAT 1086
Qy 1015 TTTAGCACTTCAACAGATGAGTGCCTACTAGGCATATGACTTACTGCGGGGGGACACA 1074
Db 1087 TACAGTGCAATCAAGCGTGTGGAGTAGCACTCAACATATGAATTAATGGGTGGGACATAGG 1146
Qy 1075 ATTCAATCTCGGCCAATAGGAGCGGATTAATACTCTCAACGCAATGGGTCTACCA--AT 1131
Db 1147 CTTAATCTCGGCCAATAGGAGGACATTAATACTCTCAACAGGACTTACTTAATAT 1206
Qy 1132 ACTCTTAATTAATCTGTAAGATTATCACTCTCTCGAGAGCTATATGGACTGAATCA 1191
Db 1207 ACTTCAATTAATCTGTAACATTACAGTTTACGTTCTCGAGAGGTTTATAGAACAGATCA 1266
Qy 1192 TATGAGAGTGCTCTATAGGGAATTTACCTTTGAACCTATTCATGGGTGCCCTACTGTT 1251
Db 1267 AATGAGGAGCAAAATAT-----ACTATTTACTCTCTGTGAAATGGAGTACCTTGGGCT 1320
Qy 1252 AGATTTAATTTTAGAACCTCCAGAACTCTTTTGAAGAGGTACTGCTAACTATAGTCAA 1311
Db 1321 AGATTTAATTTTATAAACCTCCAGAAATATTATGAAGAGGGCCCACTACCTACAGTCAA 1380
Qy 1312 CCCTATGAGTCACTCGGGCTTCAATTAAGAAATTCAGAACTGAAATACCAACAGAAACA 1371
Db 1381 CGGTATCAGGAGTTGGGATTCATTAATTTGATTGAGAACTGAAATACCAACAGAAACA 1440
Qy 1372 ACAGAACGACCAATTAATGATCATATAGTCAATAGGTTATCTCAATAGGGCTCAATTTCA 1431
Db 1441 ACAGAACGACCAATTAATGATCATATAGTCAATAGTATCTCATATAGGACTAATCATAT 1500
Qy 1432 CAATCTAGGGTGCATGACAGTATCTTCTGGAGCGACCGTAGTGCAGATCGTACAAAT 1491
Db 1501 GGAACACCTTGAGAGCACCACTTCTTCTGGAGCGATCGTAGTGCAGATCGTACGAAT 1560
Qy 1492 ACCATTAGTTCAGATAGCATAAACAAATACCATTTGGTAAATCAATCAACCTTAAATTTCA 1551
Db 1561 ACGATTGGACCAATATAGAAATACACAATACCATTTGGTAAAGCACTGAATCTTCAATTTCA 1620
Qy 1552 GGTACTCTGTAGTCAAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAATCAACGTT 1611
Db 1621 GGTGTACTGTGTGGAGGGCCAGGATTTACAGGTGGGGATATCTTCTGTAGAACAAAT 1680
Qy 1612 AATGCTAGTGTACTAAGTATGGTCTTAATTTTATATATACATCAATACAGGGTATCGC 1671
Db 1681 ACGGTGATTTGGAGATATACGATTAATATTAATGTGCCATTAATCCCAAGATATCGC 1740
Qy 1672 GTGAGAGTTCTGTATGCTCTCTCAACAAATGGTCTCGAGGGTAACTGTCCGAGGGAGT 1731
Db 1741 GTAAGATTTGGTTATGCTTCTACTACAGATTACAGATTTTCAAGGAAATTAATGGAACC 1800
Qy 1732 ACTACTTTTGATCAAGGATTCCTTAGTACTATGAGTGCAATAGAGTCTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATGGTAAATTTCTCAAGAACTATGAATAGGGGGAATAATTTAGAAATATAGA 1860
Qy 1792 TCATTTAGATTTGCAGAAATTTCTCTAGGTATAGTGCAATCGGAGTCAAACTGCTGGA 1851
Db 1861 AGTTTTGAACCTGCGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCCAAGCAATTC 1920
Qy 1852 ATAAGTAAAGTAATAATGAGGATAGACAAAGCTTTCACTTTGATAAAATTTGAATTCATT 1911
Db 1921 ACATTTGGGTCTCAGAGTTTTTCAATCAGGAAGTTTATATAGATAGAGTTCGAATTTGTT 1980
Qy 1912 CCAATTACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCCAGAGCGCGTGAAT 1971
Db 1981 CCAGCAGAGGTAACATTTTGAGGAGCAATATGATTTAGAAAGAGCACAAGGCGCGTGAAT 2040
Qy 1972 GCTCTGTTTACTAATACGAATCCAGAGATTTGAAACAGATGTGCAGATATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGNAGTTGAAACAGATGTGCAGAGTTATCATATT 2100

Qy 2032 GATCAAGTATCCAATTTTAGTGGCGTGTATTATCGGATGAATTTCTGTTAGATGAAAGAGA 2091
Db 2101 GACCAAGTGTCCAATATGGTGGCATGTTTATCAGATGAATTTTCTGTTGATGAGAAGCGA 2160
Qy 2092 GAAATTACTTGAGAAAGTGAAATATGCGAAACGACTCAGTGATGAAAGAACTTACTCCAA 2151
Db 2161 GAAATTTTGAAGAGTGAATATGCGAAGCGACTCAGTGATGAAAGAACTTACTCCAA 2220
Qy 2152 GATCAAACTTTCACATCCATCAATAAGCAACAGACTTTCATATCTACTTAATGAGCAATCG 2211
Db 2221 GATCAAACTTTCACATTCATCAGTGGCAATTAAGTTTCGCATCCATCGATGGAATCA 2280
Qy 2212 AATTTTCACATCTATCCATGAACAATCTGAAACATGAGTGGTGGGAAAGTGAGAACTTACA 2271
Db 2281 AACTTCCCTCTCTAATTAAGCTATCTGAAACATGAGTGGTGGGAAAGTGCGAATGTATACC 2340
Qy 2272 ATCAGAGAGGAAATGAGCTATTTTAAAGAGAAATTAACGTCAACTACCCGGGACTTTTAAAT 2331
Db 2341 ATTCAAGAGGAAATGAGCTATTTTAAAGAGAAATTAACGTCAACTACCCGGGACTTTTAAAT 2400
Qy 2332 GAGTGTATCCGAGCTATTTTATATCAAAATATAGAGAGTGGAAATTAAGAGCTTATACT 2391
Db 2401 GAGTGTATCCAAATTTATTTATCAAAAAATAGAGAGTCAAGATTTAAAGCTTATACG 2460
Qy 2392 CGCTACCAATTAAGAGGGTATATTTGAAGATAGTCAAGATTTTAGAGATATATTTGATTCGT 2451
Db 2461 CGCTATCAATTAAGAGGGTATATTTGAAGATAGTCAAGATTTTAGAGATATATTTGATTCGT 2520
Qy 2452 TATTAATGGGAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGATGCGCGCTTTCA 2511
Db 2521 TACAATGCAAGCATGAAACATTTGGATGTTCCAGGTACCGATTTCCCTATGCGCGCTTTCA 2580
Qy 2512 GTTCAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGACACACATTTTGAATGG 2571
Db 2581 GTTGAAGCCCAATCGGAAGTGGGAGAACCAATCGATCGGACACACATTTTGAATGG 2640
Qy 2572 AATCCTGATCTAGATTTGTTCTCGAGAGATGGAGAAATTTGGCGCATCATTTCCCATCAT 2631
Db 2641 AATCCTGATCTAGATTTGTTCTCGAGAGATGGAGAGATGTCGCGCATCATTTCCCATCAT 2700
Qy 2632 TTCTCTTTGGATATGATATTTGGATGCAACAGACTTGCATGAGAAATCTTAGGGGTGTGGTG 2691
Db 2701 TTCACTTTGGATATGATTTGGTGTGCAACAGACTTGCATGAGAACTTAGGGGTGTGGTG 2760
Qy 2692 GTATTTCAAGATTAAGACGAGAGAGTCACTCAAGACTTAGGGAATCTCGGAATTTATTTCAA 2751
Db 2761 GTATTTCAAGATTAAGACGAGAGAGTCACTCAAGATTTAGGAATTTAGGAATTTATCGAA 2820
Qy 2752 GAGAAACCATTTATTTAGGAGAGCACTGTCTCGTGTGAAGAGAGAGAGAAATTTGGAGA 2811
Db 2821 GAGAAACCATTTATTTGGAGAGCACTGTCTCGTGTGAAGAGAGAGAGAAATTTGGAGA 2880
Qy 2812 GACAAACGTGAAACATCAATTTGAAACAAACAGAGTATATACAGAGGCAAAAGAGCT 2871
Db 2881 GACAAACGTGAAACATCAATTTGAAACAAACAGAGTATATACAGAGGCAAAAGAGCT 2940
Qy 2872 GTGATGCTTTTATTTGTAGATTTCTCAATATATAGATTTTACAGCGGATACAAACATTTGGC 2931
Db 2941 GTGATGCTTTTATTTGTAGATTTCTCAATATATAGTCAATTTTACAGCGGATACAAACATTTGGC 3000
Qy 2932 ATGATTCATCGGAGAGATAAATTTGTTTCATCGAATTTGAGAGGCTTATCTCTCAGAAATTA 2991
Db 3001 ATGATTCATCGGAGAGATAAATTTGTTTCATCGAATTTGAGAGGCTTATCTTTTCAGAAATTA 3060
Qy 2992 TCTGTTATCCCGGTGTAATGCGGAAATTTTGAAGATTTAGAAAGTTCGCAATTTATCACT 3051
Db 3061 CCTGTTATCCCGGTGTAATGCGGAAATTTTGAAGATTTAGAAAGTTCGCAATTTATCACT 3120
Qy 3052 GCAATCTCCCTATACGATGCGGAAATGTGTTTAAATTTGATTTTAAATTTGATTTTAAATTTGATTTA 3111
Db 3121 GCAATGCTCTTATACGATGCGGAAATGTGTTTAAATTTGATTTTAAATTTGATTTTAAATTTGATTTA 3180

3112 GCATGCTGGAAATGTAAGAGGCGCATGTAGATGTACAAACAGAGCCATCACCGTTCTGTCTCTT 3171
3181 ACATGTTGGAATGTAAGAGGCGCATGTAGATGTACAAACAGAGCCATCATCGTTCTGACCTT 3240
3172 GTTATCCAGAAATGGAGAGGAGAGTGTACAGAGAGTGTGCGTCTGTCGCGGGCGGTGGC 3231
3241 GTTATCCAGAAATGGAGAGGAGAGTGTACAGAGAGTGTGCGTCTGTCGCGGGCGGTGGC 3300
3232 TATATCTCTCGTGTACAGAGGTAACAAAGAGGAGATATGGAGAGGTTGTGTACAGATCCAT 3291
3301 TATATCTCTCGTGTACAGAGGTAACAAAGAGGAGATATGGAGAGGTTGTGTACAGATCCAT 3360
3292 GAAATCGAGAACATACAGAGCACTAAATTTAAAACTGTGTGAAGAGAGAGTGTAT 3351
3361 GAAATCGAGAACATACAGAGCACTAAATTTAAAACTGTGTGAAGAGAGAGTGTAT 3420
3352 CCAACGATACAGAGAGTGTATGATATATCTGACACCAAGTGTACAGAGTGTATGATAT 3411
3421 CCAACGATACAGAGAGTGTATGATATATCTGACACCAAGTGTACAGAGTGTATGATAT 3480
3412 TCCCGTAAATGCTGATATGAGGATGCAATATGAAGTTGATATACAGATCTGTAAATTAC 3471
3481 TCCCGTAAATGCTGATATGAGGATGCAATATGAAGTTGATATACAGATCTGTAAATTAC 3540
3472 AAACCGATATGAAGAGAGAAAGTGTATACAGATGTACAGAGATTAATCAATGTGATAT 3531
3541 AAACCGATATGAAGAGAGAAAGTGTATACAGATGTACAGAGATTAATCAATGTGATAT 3600
3532 GACAGAGGATGTGAATTTATCCACCACTACAGAGTGTATATGACAAAGAAATTAGAA 3591
3601 GACAGAGGATGTGAATTTATCCACCACTACAGAGTGTATATGACAAAGAAATTAGAA 3660
3592 TACTTCCAGAGAAACCGATATGAGTATGAGATTTGAGATTTGAGAGAAACCGAGAGGATTTAT 3651
3661 TACTTCCAGAGAAACCGATATGAGTATGAGATTTGAGATTTGAGAGAAACCGAGAGGATTTAT 3720
3652 GTAGACAGCGTGAATTTACTCTTATCGAGAGATAG 3687
3721 GTAGATAGCGTGAATTTACTCTCTATCGAGAGATAG 3756

RESULT 11
US-08-040-751-4
; Sequence 4, Application US/08040751
; Patent No. 5407825
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M
; APPLICANT: Sick, August J
; TITLE OF INVENTION: No. 5407825el Bacillus thuringiensis isolates
; TITLE OF INVENTION: active against Lepidopteran Pests and Genes Encoding No. 5407825
; TITLE OF INVENTION: Lepidopteran-active toxins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID R. SALIWANCHIK
; STREET: 2421 N.W. 41st STREET, SUITE A-1
; CITY: GAINESVILLE
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/040,751
; FILING DATE: 19930329
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SALIWANCHIK, DAVID R.
; REFERENCE/DOCKET NUMBER: MA39.C1.D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800
TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3522 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick
; CLONE: 81A2
; US-08-040-751-4
Query Match 48.6%; Score 1793.4; DB 1; Length 3522;
Best Local Similarity 71.9%; Pred.No. 0;
Matches 2497; Conservative 0; Mismatches 861; Indels 117; Gaps 7;
QY 228 ACAATAGCTAGTCTTTTATAGTCTTTTGTGGTGAATTTATGGCCCGCGGAGAGATCA 287
Db 147 ACTAGGGGATTTTATCTTGGCTTGTGATGTAATATGGGGGCTATAGGTCCTTCACA 206
QY 288 GTGGGAAATTTTCTTAGAACATGTGCAACAATTTAATTAATCAACAATAACAGAAATGC 347
Db 207 ATGGGATATATTTTAGAGCAAAATTTGAGCTATTTGATCGGCCCAAGAAATAGAGGAATTCG 266
QY 348 TAGGAATACGGCACTTCTCGATTTACAGGTTTACAGAGATCTCTTTAGAGCCTATCAACA 407
Db 267 TAGGAATACGGCACTTCTCGATTTACAGGTTTACAGGTTTACAGGTTTACAGGTTTAC 326
QY 408 GTCACTTGAAGATTTGGCTAGAGAAACCGTGTATGATGATGATGATGATGATGATGATGAT 467
Db 327 TCGTTTAAATAACTGGGAGTAGATCTTAAATCCAGCATTAAGAGAAGAGATGCGTAT 386
QY 468 CCAATATATAGCTTTAGAACCTTGTCTTAAATCGGATGCGCTTTTCGCAATTTAGAA 527
Db 387 TCAATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
QY 528 CCAAGAAATTTTCAATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 587
Db 447 TTATGAATTTCTCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 506
QY 588 GAGAGATGCTCTCTTTTGTGGTGAATTTGGGCTTACATCGAGGAAATTTCAAGTTA 647
Db 507 GAGAGATGTTTCAGTGTGTTGGACAACTTTGGGATTTGATGATGATGATGATGATGATGAT 566
QY 648 TTATGAGCGCAAGTGTGAACAAACGAGAGATTTTCCGACTATTTGGTGAATTTAGTAA 707
Db 567 TTATAATGATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
QY 708 TACAGCTCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764
Db 627 TACGGGTTAATCGTTTACCGATTAATGAAGGGTACGAGGATGGCAAGATTTAATAG 686
QY 765 ATTCCGTAGAGATCTAACTGTTGGGTTATAGTCTAGTGGCACTATTTCCAGCTATGA 824
Db 687 GTTTAGAGAGAGTAAACAATATCAGTATTTAGATATTTCTTTTCTTTTCTTTTCTTTTCT 746
QY 825 CACTCGCACTTATCCCAATTAATACAGAGTGTCTAGTGTAAACAAGGAAAGTTTATACAG 884
Db 747 TTCTAGATTTATCCCAATTTCCGCAATCTCTCAATTAACCGGGAAGTATATACAGATCC 806
QY 885 AATTGGAGCAACAGGGGTAAATATGCAAGTATGAATTTGGTATTAATTAATGACCTTC 944
Db 807 -----GGTAAATTAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
QY 945 GTTTTCCGCTATAGAGACTCGGTTATCCGAAGCCCGCATCTACTTGATTTTCTAGAAC 1004

Db	843	TTTCGAGGATATGAAATTCAGCTATTAGAGTCCCATCTTATGAGATTTCTTAATAAA	902
Qy	1005	ACTTTACAATTTTTTAGCAGCTTCATCACGATGGAGTGTCTACTAGGCATATGACTTACTGGCG	1064
Db	903	TATAATATTGACACTGATTATA-----ATTAGAGCGGTCCTACTATTGGGC	947
Qy	1065	GGGGCACAATATTCOAATCTCGGCCCAATAGAGAGGGGATTAATACTCTCAACGATGGGTC	1124
Db	948	GGGGCATCGTGTAACTTCTCATTTTACCGGTAGTTCGCAAGTGATAGCTCCCTCAATA	1007
Qy	1125	TACCAATACTTCTATTAACTCCTGTAAAGTTATCATTTCTTCTCTCGAGACGTATTTGGAC	1184
Db	1008	CGGGATTAACATGCAAAACGAGAACCGAGTGCAGACTATTGCTCCTAGCAGCTTTTCCAGGTCT	1067
Qy	1185	TGAATCATATGCGAGGTGCTTCTATGGGGAATTTACCTTTGAACTTATCATGGTGTCCC	1244
Db	1068	TAACTCTATTATAGACACATCATCAGACCCTTCTTCGGAAGATCCGATATATTATGCC	1127
Qy	1245	TACTGTTAGATTAAATTTTATAGAACCCCTCAGAAATATCTTTTGAAGAGGTACTGCTAACTA	1304
Db	1128	AACATTAGGAATAAATCTAGTGCAGGGGGTAGGATTCAATTCAACCAAAATATGGTGAAGT	1187
Qy	1305	TAGTCAACCCATATGAGTCACTGGGCTTCAATTTAAAGATTTCAGAACTTCAGAAATTCACAC	1364
Db	1188	TCTATATAGAAGGAGAGGAACGATAGATTCTCTTGATGAGTTGCCAATTGA-----CGG	1241
Qy	1365	AGAAACAACAGAACGACCAAAATTATGAATCATATAGTCATAGGTATCTCACATAGGGCT	1424
Db	1242	TGAGAAATTCATTAGTTGGATATAGTCATAGATTAAAGTCAGCTTACATTAACAGAGTCGTT	1301
Qy	1425	CATTTCAAATCTTAGGTGCAATGACAGATATATCTTTGGACGCACCGTAGTCAGATCG	1484
Db	1302	ATATAACTATAATAACTAGCTTGCCAACATTTGTTTGGACACATCACAGTGTCTACTGA	1361
Qy	1485	TACAAATACCAATTAGTTTCAGATAGCATAAACAAATACCATTGGTAAATCATTTCAACCT	1544
Db	1362	TCGAAATATAATCTATCCGATGTGAATTCACAAATACCAATCGGTAAATCATTTCCCT	1421
Qy	1545	TAATTCAGGTACCTCTGTAGTTCAGTGCCCAGGATTTTACAGGAGGGGATATAATCCGAAC	1604
Db	1422	TACTTCAGGTACCTCTGTAGTCAGAGGCCACAGATTTTACAGGAGGGGATATCATCCGAAC	1481
Qy	1605	TAACTGTTAATGGTAGTGTACTAAATATGGGTCTTAATTTTAAATAATCATTCATTACAGCG	1664
Db	1482	TACGTTAAATGGTAAATGCTACTAAGTATGAGTCTTAATTTTATAGTAATACATCATTCAGCG	1541
Qy	1665	GTATCGCGTGAGAGTTGCTTATGCTGTCTTCTCAAAACATGGTCTCTGAGGGTAACTGTCCG	1724
Db	1542	GTATCGCGTGAGAGTTGCTTATGCTGTCTTCTCAAAACAAATGGTCAATGAGAGTAAATGTTGG	1601
Qy	1725	AGGAGTACTACTTTTTCATCAAGGATTCCTGTAGTACTATGAGTGCAAAATGAGTCTTTGAC	1784
Db	1602	AGGAGTACTACTTTTGTATCAAGATTTCCCTGTAGTACTATGAGTGCAAAATGGTCTTTGAC	1661
Qy	1785	ATCTCAATCATTTAGATTTCAGAAATTTCTGTAGGTATTAGTGCATCTGGCAGTCAAAAC	1844
Db	1662	ATCTCAATCATTTAGATTTCAGAAATTTCTGTAGGCAATTTAGTACATCTGGCAGTCAAAAC	1721
Qy	1845	TGCTGGAAATAGTATTAAGTAAATATGAGGTAGACAAAACGTTTTTCACTTTGATATAAATTTGA	1904
Db	1722	TGCTGGAAATAGTATTAAGTAAATATCAGGTAGACAAAACGTTTTTCACTTTAGATAGAATTTGA	1781
Qy	1905	ATTCAATTCCAATTTACTCMAACCTTCGAAGCAGAAATACGATTTAGAAGGGCGCAGGGC	1964
Db	1782	ATTTATCCCCAGTTGATGCAACATTTGAAGCAGAAATATGAATTTAGAAGAGCACAAAAGGC	1841
Qy	1965	GGTGAATGCTCTGTTTACTAATACGAATCCAGAAAGATTTGAAAACAGATGTGACAGATTA	2024
Db	1842	GGTGAATTCGCTGTTTACTTCTTCCATCAATTCGAGTTTAAACAGATGTGACGATTA	1901
Qy	2025	TCATATTGATCAAGTATCCAATTTAGTGGCGGTGTTATTCGGATGAATTTCTGCTTAGATGA	2084

Db	1902	TCATATTGGATCAAGATATCCAAATTTAGTGTAGAAATTTGTTTATCCGATGAAATTTTGTCTCGTGGANGA	1961
Qy	2085	AAAGAGAGAATTTACTTTCGAGAAAGTCAAATATATGCGAAACGACTCAGTGCATGAAAGAAAATCTT	2144
Db	1962	AAACGGAGAATTTGTCGAGAAAGTCAAAACATGCGAAGGACTCAGTGCATGAGCGGAATTT	2021
Qy	2145	ACTCCAAGATCCAAACTTTCATCATCCATCAATTAAGCAACCGAGCTTCAATCTCTACTAATGA	2204
Db	2022	ACTTCAAGATCCAAACTTTCAGAGGGATCAATAGCAACACGAC-----	2064
Qy	2205	GCAATCGAATTTTCACATCTATCCATGAAACAACTCTGAAACATGGATGGTGGGGAAGTGAAG	2264
Db	2065	-----CGTGGCTGGAGAGGAAAGTACGGA	2087
Qy	2265	CATTACAATCCAGGAAGAAATGACGTATTTAAAGAGAAATTTACGTGCACACTACCGGGGAC	2324
Db	2088	TATTTACCATCCAAGGAGGAGTGCCTATTTCAAGAGAAATTCGTGCACATCCAGGTAC	2147
Qy	2325	TTTTTAATGAGTGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAAAGC	2384
Db	2148	CTTTTGATGAGTCTATCCAAACGTATTTGTATCAAAAAATAGATGAGTCGAAAAATTTAAAGC	2207
Qy	2385	TTTACTCGCTACCAATTAAGGGGTATATTGAGATAGTTCAGATTTAGAGATATATTT	2444
Db	2208	CTATAACCGTGTACCAATTTAAGAGGGTATATCGAAGATAGTCAAGACTTTGAAATCTCTATTT	2267
Qy	2445	GATTCGTATTAATCGCGAAACATGAAACATTTGGATTTTCCAGGTACCGGTCCGTATGCGC	2504
Db	2268	AAATTCGTACATGCAAAACAGNAACAGTAATGTACAGAGTACGGGTTCTTATGCGC	2327
Qy	2505	GCTTTTCAGTTGAAAGCCCAATCCGAAGGTCGGAGAAACCGAATCGATGCGCACCAATTT	2564
Db	2328	GCCTTCAGTCGAAAGTCCAATTTGGAAGGTGTGGAAGACCGAATCCGGTGTGTGCCACACCT	2387
Qy	2565	TGAATGGAATCCTGATCTAGATTTGTTTCCCTCGAGAGTGGAGAAATATGTCCGCATCATTC	2624
Db	2388	TGAATGGAATCCTGATTTAGATTTGTTTCCCTCGAGAGCGGGGNAAAATGTGCACATCATTC	2447
Qy	2625	CCATCATTTCTCTTTGGATATTTGATATTGGATGCAAGACTTTCATGAGAGAAATCTTAGGCGT	2684
Db	2448	CCATCATTTCTCTTTGACATTTGATTTGGATGCAACAGACTTTCGAGAGAGATCTAGGCGT	2507
Qy	2685	GTGGGTGGTATTTCAAGATTAAAGACGCAAGGAGGTCTGCAAGACTAGGGGAATCTGGAATTT	2744
Db	2508	GTGGGTGGTATTTCAAGATTAAAGACGCAAGGAGGTATTGCAAGATTAGGAATCTGGAATTT	2567
Qy	2745	TATTTGAAGAGAAACCAATTTAGGAGAGACATCTCTCGTGTGAAGAGACGAGAAAAA	2804
Db	2568	TATCGAAGAGAAACCAATTAATGGAAGACATCTCTCGTGTGAAGAGACGCGGAAAAAA	2627
Qy	2805	ATGGAGAGACAAACGTGAAAACTACAAATTTGGAACAAAAACAGTATATACAGAGGCAAA	2864
Db	2628	ATGGAGAGACAAACGGGAAANACTACAAATTTGGAACAAAAACAGTATATACAGAGGCAAA	2687
Qy	2865	AGAAGCTGTGGATGCTTTTATTTGTAGATTTCTCAATATAATAGATTTCAACGGGATACAAA	2924
Db	2688	AGAAGCTGTGGATGCTTTTATTTGTAGATTTCTCAATATGATAGATTTCAACGAGATACAAA	2747
Qy	2925	CATTGGCATGATTCATCGCGCAGATAAACCTTGTTCATTCGAATTTGAGAGGCTTATCTGTC	2984
Db	2748	CATTGGTATGATTCATCGCGCAGATGACTTGTTCATTCAGATCCACGAGGCTTATCTCTCC	2807
Qy	2985	AGAATTATCTGTTATCCCGGTGTAAATCGGAAAAATTTTGAAGAAATTAAGAAGTCGCAT	3044
Db	2808	AGAACTACCTTTTCAATCCAGAAATAAATGTGGTGAATTTTGAAGAAATTAAGAAACCGTAT	2867
Qy	3045	TATCACTGCAATCTCCCTATACGATCGGAAAAATGTCTGTTTAAAAATTTGGTATTTTAATAA	3104
Db	2868	TTCTACTGCAATATCCCTATATGATCGGAAATGTTCATTTAAAAATGGCGATTTCAATAA	2927
Qy	3105	TGGATTAGCATGCTGGAATCTAAAGGGCATGTAGAT---GTACAAACAGAGCCATCACCG	3161
Db	2928	TGGCTTATCATGTGTGAAACGTGAAAGGGCATGTAGATGTAGTAGAACAACCAACACCG	2987

Db 843 TTTTCGAGAGTATTGAAATTCAGCTATTAGAAAGTCCCATCTTATGGATTTCTTAAATAA 902
Qy 1005 ACTTACAAATTTTAGCATTCTATCAGATGGAGTCTACTAGGCATATGACTTACTGCGG 1064
Db 903 TATAATTTATGACCTGATTTA-----ATTAGGCGTTTCACTATTGGGC 947
Qy 1065 GGGGCACACAAATTCATCTCGCCCAATAGGAGCGGATTAATACTCAACGCGCATGGGTC 1124
Db 948 GGGGCATCGTGTAACCTTCTCATTTTACCGGTAGTTGCGAAGTGATAAGCTCCCTCAATA 1007
Qy 1125 TACCAATACTTCTATTAATCTCTGTAAGATTAATCATTTCTCTCGAGACGTAATATGGAC 1184
Db 1008 CGGGATAACTGCAACCGCAGAACCGAGTGAATTTGCTCTCTAGCACTTTTCCAGGTCT 1067
Qy 1185 TGAATCATATGACGAGGTCTTCTATGGGAATTTACCTTGAACCTATTTCTATGGTGTCC 1244
Db 1068 TAATCTATTTATAGAACACTATCAGACCCCTTTCTTCGAGAGTCCGATAATATTATGCC 1127
Qy 1245 TACTGTTAGAATTTAATTTTAGGAACCTCAGAAATACTTTTGAAGAGGTACTGTCTAACTA 1304
Db 1128 AACATTAGGAATAATATGATGTCAGGGGTAGGATTCATTCAACCAATATATGGTGAAGT 1187
Qy 1305 TAGTCAACCCWTAGAGTCACTGGGGCTCAATTAAGAAATTCAGAAACTGAATTAACCACC 1364
Db 1188 TCTATATAGAAGGAGGAAACAGTAGATTTCTCTGATGAGTTGCCAATTGA-----CGG 1241
Qy 1365 AGAAACAACAGAACCAACCAATTAATGAATCATATAGTCAATAGTATTCTCATAGGCGT 1424
Db 1242 TGAGAAATCATTTAGTTGGATATAGTCATAGATTAAGTCACTGTTCAATTAACCAAGTCGTT 1301
Qy 1425 CATTTTCACAATCTAGGGTGCATGTACCACTATATTCTTGGACGACCTAGTGCAGATCG 1484
Db 1302 ATATAACTAATATAAATAGCTTGGCAACATTTGTTTGGACACATCACAGTGTACTGA 1361
Qy 1485 TACAAATACCAATGTTTCAGATAGCATAAACAAATACATTTGGTAAATCATTTCAACCT 1544
Db 1362 TCGAAATATAATCTATCCGGATGTAATTAACAAATACCAATTTGGTAAATCATTTCCCT 1421
Qy 1545 TAAATTCAGTACTCTGTAGTCAGTGGCCCGAGGATTTACGAGGGGATATAATCCGAAC 1604
Db 1422 TACTTCAGGTACTCTGTAGTCAGAGGCCAGGATTTTACAGAGGGGATATCATCCGAAC 1481
Qy 1605 TAACTGTAATGTTAGTGTACTAAGTATGGGTCTTAATTTTAATAATACATCATTTACAGCG 1664
Db 1482 TAACTTAAATGTAATGTTACTAAGTATGAGTCTTAATTTTAGTAATAATCATTTACAGCG 1541
Qy 1665 GTATCGGTGAGAGTTGTTATGCTGTTCTCAAAATGCTCTGAGGGTAACTGTGCG 1724
Db 1542 GTATCGGTGAGAGTTGTTATGCTGCTTCTCAAAATGCTCTCATGAGTAAATGTTGG 1601
Qy 1725 AGGGAGTACTACTTTTGTATCAAGGATTCCTAGTACTATGATGCAATGATCTTTGAC 1784
Db 1602 AGGGAGTACTACTTTTGTATCAAGGATTCCTAGTACTATGATGCAATGCTTTGAC 1661
Qy 1785 ATCTCAATCATTTAGATTTGAGATTTCTGTAGGTATTAGTGCATCTGGCAGTCAAAC 1844
Db 1662 ATCTCAATCATTTAGATTTGCGAATTTCTGTAGGCAATAGTACATCTGGCAGTCAAAC 1721
Qy 1845 TGCTGGAATAAGTATAAGTAAATGTCAGGTAGACAAACGTTTCACTTTGATAAAATGGA 1904
Db 1722 TGCTGGAATAAGTATAAGTAAATCCAGGTAGACAAACGTTTCACTTTAGATGAATTTGA 1781
Qy 1905 ATTCATTCCAATTAATCTGCAACCTTCGAGCAGCAATACGATTTAGAAAGGCGCAGAGCG 1964
Db 1782 ATTTATCCAGGTGTATGCAACATTTGAAAGCAGCAATATGATTTAGAAAGCAGCAAAAGCG 1841
Qy 1965 GGTGAATGCTCTGTTTACTTAATACCAATCCCAAGAGATTTGAAACAGATGTGACAGATTA 2024
Db 1842 GGTGAATTTGCTGTTTACTTCTTCCAAATCAATCAAGTTTAAACAGATGTGACGGATTA 1901
Qy 2025 TCATATTGATCAAGTATCAATTTTATAGTGGCGTGTATTTCGGATGAATTTCTGCTTAGATGA 2084

Db 1902 TCATATTGATCAAGTATCCAATTTTAGTAGATTGTTTATCCGATGAATTTTGTCTGGATGA 1961
Qy 2085 AAAGAGAAATTTACTTGGAAAGTGAATATGCGAAACGACTCAGTGATGAAGAAACTT 2144
Db 1962 AAAGCGAATTTGTTCGAGAAAGTCAACACTGCGAAGCGACTCAGTGATGAGCGGAATTT 2021
Qy 2145 ACTCAAGATCCAACTTTCACTCATCATCAATTAAGCAACAGACTTTTATCTACTTAATGA 2204
Db 2022 ACTTCAAGATCCAACTTTCAAGGGATCAATAGCAACCCAGAC----- 2064
Qy 2205 GCAATCGAATTTTCACTATCTATCCATGAACAATCTGAAACATGATGGTGGGGAAGTGAGAA 2264
Db 2065 -----CGTGGCTGGAGAGGAAGTACGGA 2087
Qy 2265 CATTACAATCCAGGAAGGAAATGACGTATTTAAAGAGAAATTAACGTCACTACACCGGGAC 2324
Db 2088 TATTACCATCCAGAGGAGATGACGTATTTCAAAGAGAAATTAACGTCACTACACCGAGTAC 2147
Qy 2325 TTTTAATGAGTGTATTCCGAGCTATTTATATCAAAAAATAGGAGAGTCCGAATTTAAAAAGC 2384
Db 2148 CTTTGTAGTGTCTATCCAACTGATTTGTATCAAAAAATAGATGAGTCCGAATTTAAAAAGC 2207
Qy 2385 TTATACCTGCTACCAATTAAGAGGGTATTTAGAGATAGTCAAGATTTTAGAGATATATTT 2444
Db 2208 CTATAACCGTTTACCAATTTAAGAGGGTATTCGAAGATAGTCAAGACTTTAGAAATCTATTT 2267
Qy 2445 GATTCGTTATTAATCGGAAACATGAACATTTGGATGTTTCCAGGTACCGAGTCCGATATGCC 2504
Db 2268 AATTCGCTACAAATGCAAAACACGAAACAGTAATTTGTACAGGTACGGGTTCTTTATGCC 2327
Qy 2505 GCTTTCAGTTGAAAGCCCAATCCGAAGGTGCGGAGAACCGAATCGATCGGCACCACTTT 2564
Db 2328 GCTTTCAGTCCAAAGTCCAATTTGGAAGGTGGAAGACCGAATCGGTGTGTGCCACACT 2387
Qy 2565 TGAATGGAATCCTGATCTAGATGTTCTCTGAGAGATGGAAGAAATGTGCGCATCATTC 2624
Db 2388 TGAATGGAATCCTGATTTAGATTTGTTCTGAGAGACGCGGGAAGAAATGTGCACACTATTC 2447
Qy 2625 CCATCATTTCTTTGGATATTGATATTGGATGCACAGACTTGCATGAGATCTAGGGGT 2684
Db 2448 CCATCATTTCTTTGGACATTTGTTGGATGCACAGACTTGAAGAGGATCTAGGGGT 2507
Qy 2685 GTGGGTGATTTCAAGATTTAAGACGCGAGGAAGGTCTGCAAGACTTAGGGAATCTGGAAT 2744
Db 2508 GTGGGTGATTTCAAGATTTAAGACGCGAGGAAGTTATGCAAGATTTAGGAATCTGGAAT 2567
Qy 2745 TATTGAAGAGAAACCAATTTATAGAGAGCACTGTCTGCTGTGAGAGAGAGAGAGAAAA 2804
Db 2568 TATCGAAGAGAAACCAATTTAATTTGGAAGCACTGTCTGCTGGAAGAGAGCGGAAAAA 2627
Qy 2805 ATGAGAGACAAACGTCGAAACCTACAAATTTGGAACCAAAACGATATATACAGAGGCAAA 2864
Db 2628 ATGAGAGACAAACGCGGAAAAACCTACAAATTTGGAACCAAAACGATATATACAGAGGCAAA 2687
Qy 2865 AGAAGCTGTGATGCTTTTATTTGTAGATTCTCAATTAATAGATTACAAGCGGATACAAA 2924
Db 2688 AGAAGCTGTGATGCTTTTATTTGTAGATTCTCAATATGATAGATTACAAGCAGATACAA 2747
Qy 2925 CATTTGCAATGATTCATGCGCAGATAAACTTTGTTTCTGGAATTCGAGAGGCTTATCTGTC 2984
Db 2748 CATTTGATGATTCATGCGCAGATAGACTTTGTTTCTCATCAGATCCACAGGCTTATCTTCC 2807
Qy 2985 AGAATTAATCTGTTATCCGGGTGTAATGCGGAATTTTGAAGATTTTAGAGGTCGAT 3044
Db 2808 AGAATTAATCTGTTATCCAGGAATAAATGTTGGTGAATTTTGAAGATTTAGAAAAACCGGAT 2867
Qy 3045 TATCACTGCAATCTCCCTTATACGATGCGAGAAATGTCGTTAAAAATGCTGATTTAATAA 3104
Db 2868 TTCTACTGCAATATCCCTATATATGATGCGAGAAATGTCATTTAAATGCGGATTTCAATA 2927
Qy 3105 TGGATTAAGATGCTGGAATGTAAGAGGCAATGTTAGAT---GTAACAAGAGCCATCACCG 3161
Db 2928 TGGCTTATCATGCTGGAACGTTGAAAGGCGATGTAGATGTAGTAAGACAAACCAACCCG 2987

QY 3162 TTCTGTCCTTGTATCCAGATGGGAGCAGAGTGTCAAGAGTTCGCGTCTGTCC 3221
Db 2988 TTCGGTCTTGTGTCGGAATGGGAGCAGAGTGTCAAAACAATTCGTGTCGTCC 3047
QY 3222 GGGCGTGGGTATATCTCCGTGTACAGAGGTCAAGAGGGATATGGAGAGGTTGTGT 3281
Db 3048 GGGCGTGGGTATATCTCCGTGTACAGGTCAAGAGGGATATGGAGAGGTTGCGT 3107
QY 3282 AACGATCCATGAATCCAGAACATACAGACGAACCTAAATTTAAAACTGTGAAGA 3341
Db 3108 AACCATCATGAGATCGAGAACAAATACAGACGAACCTAAATTTAAAACTGTGAAGA 3167
QY 3342 GGAAGTGTATCCACGATACAGGAAGTGTATGATTAATCTGCACACCAAGGTACGC 3401
Db 3168 GGAAGTGTATCCACGATACAGGAAGTGTATGATTAATCTGCACACCAAGGTACGC 3227
QY 3402 AG-----TATGTAATCCCGTAATGCTGGATATGAGGTGTCATATCAAGTTGA 3449
Db 3228 AGATCCACAGATCATGTATTTCCCGTAATATCAGATATGAGATGTCATATGAATGAA 3287
QY 3450 TACTACAGCATCTGTTAAATACAAACCGACTTATGAAGAAGAAACGTATACAGATGTACG 3509
Db 3288 TACTACAGCATCTGTTAAATACAAACCGACTTACGAAGAAGAAAGGTATACAGATGTACA 3347
QY 3510 AAGAGATAATCATGTGCAATATGACAGAGGGTATGTGAATTTCCACCCTACACAGTGG 3569
Db 3348 AGGAGATAATCATGTGTAATATGACAGAGGGTATGTGAATTTACGACCATACACAGTGG 3407
QY 3570 TTATATGACAAAAGAAATAGATACTTCCAGAAACCGATAAGGTATGGATTCGAGATTGG 3629
Db 3408 TTATGTGACAAAGAAATAGATACTTCCAGAAACCGATAAGGTATGGATTCGAGATTGG 3467
QY 3630 AGAAACGGAAGGAAGTTTATTTGTAGACAGCGTGGAAATTAATCTCTTATGGAGGAA 3684
Db 3468 AGAAACGGAAGGAAGTTTATTTGTAGACAAATGTCGAATTAATCTCTTATGGAGGAA 3522

RESULT 13
US-08-962-190-1
; Sequence 1, Application US/08962190
; Patent No. 5985267
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: No. 5985267el Bacillus thuringiensis Isolates
; TITLE OF INVENTION: Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/962,190
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/291,368
; FILING DATE:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA50.C1

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick
; CLONE: 81A2
; US-08-962-190-1

Query Match 48.6%; Score 1793.4; DB 2; Length 3522;
Best Local Similarity 71.9%; Pred. No. 0;
Matches 2497; Conservative 0; Mismatches 861; Indels 117; Gaps 7;
QY 228 ACAATAGCTAGTCTTTTATAGTCTTTCTGTTGGTGAATTTATGGCCCGCGGCAGAGATCA 287
Db 147 ACTAGGGGATTTTATATCTTTGGCTTGTGATTAATATGGGGGCTATAGGTCCTTCACA 206
QY 288 GTGGGAAATTTCCCTAGAACATGTGAAACAACTTATAAATCAACAATAACAGAAAATGC 347
Db 207 ATGGGATATATTTTAGACCAATTTGAGCTATTTGATCGGCCAAAGAAATAGAGAAATCGC 266
QY 348 TAGGAATACGGCACTTGTCTCGATTTACAAGSTTTTAGGAGATTCCTTTAGAGCCTTATCAACA 407
Db 267 TAGGAATACGGCAATTTTCTAGATTTACAAGGGCTTAAGCAATCTTTACCGAATTTTACACAA 326
QY 408 GTCATTGGAAGATTGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTTATAC 467
Db 327 TGCCTTTTAAAACTGGGAAGTAGATCCTACTAATCCAGCATTAAGAGAGAGATGCGTAT 386
QY 468 CCAATATATAGCTTTAGAACTTGATTTCTTAATGCGATGCGCTTTTCGCAATTAGAAA 527
Db 387 TCAATTTAATGACATGAACAGTGCCTTACACACAGCTATTCCTCTTTTTCAGTTCAGG 446
QY 528 CCAAGAGTTCCATTTAATGATGATATGCTCAAGCTGCAAAATTTACACCTATTTATT 587
Db 447 TTATGAAATTCCTCTTTTATCAGTATATGTTCAAAGCTGCAAAATTTACATTTATCGGTTT 506
QY 588 GAGAGATGCTCTCTTTTGGTGTAGTGAATTTGGGCTTACATCGCAGGAAATTTCAACGTTA 647
Db 507 GAGAGATGTTTCAGTGTTTGGACAACGTTGGGATTTGATGTAGCAACCAATCAATAGTCG 566
QY 648 TTATAGCGCCAGTGGAAACAAACGAGAGATTTATCCGACTATTTGCGTAGAATGGTATAA 707
Db 567 TTATATGATTTAAGTGTATTTAGGCTTTATTTGGCGAATATCTGATTTATGCTGTAGTTGTTATA 626
QY 708 TACAGTCTAAATAGCTTGGAGGGGCAAAATGCGG---CAAGTTGGGTGCGTTTATATCA 764
Db 627 TACGGGTTAAATCGTTTACCACGTAATGAAGGGGTACGAGGATGGCAAGATTTAATAG 686
QY 765 ATTCCGTAGAGATCTAAACGTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGA 824
Db 687 GTTTAGAAAGAGAGATTAAACCAATATCAGTATTAGATATTTCTTTTTCCTCAAAATACGA 746
QY 825 CACTCGCACTTATCCAAATAAATACAGAGTGTCTCAGTTTAAACAAAGGAAAGTTTATACAGACGC 884
Db 747 TTCTAGATTATATCCAAATTCGCGCAATCTATCAATTAACGCGGAAGTATATACAGATCC 806
QY 885 AATTGAGCAACAGGGGTAAATATGCAAGTATGAATTTGGTATTAATAATATGACCTTC 944
Db 807 -----GGTAATTAATAATACTGATTATAGAGTTATACCCCAAG 842

945 GTTTCCGCTATAGAGCTGCGGTTATCCGAAGCCGCGATCTACTGTGATTTTCTAGAAC 1004
843 TTTCGAGAGTATTGAAATTCAGCTATTAGAAAGTCCCATCTTATGGAATTTCTTAAATA 902
1005 ACTTACAAATTTTAGACCTTCATCAGCATGGAGTGTCTACTAGGCATATGACTTACTGGG 1064
903 TATAATTTATGACACTGATTTA-----ATTAGAGCGGTTCACTATTGGGC 947
1065 GGGGACACAAATCAATCTCGGCCAATAGAGCGGATTAATAATCCTCAAGCATGGGTC 1124
948 GGGGCATCGTGTAACTTCTCATTTTACCGGTAGTTCGCAAGTGAAGTCCCTCAATA 1007
1125 TACCAATATCTTATTAATCTCTGTAGATTAATCTTCTCTCGAGAGCTATATGGAC 1184
1008 CGGGATAACTGCAACCGCAAGACCGAGTCGAATATTGCTCTAGCACTTTTCCAGGTCT 1067
1185 TGAATCATATGACGAGTGTCTTATGGGAATTTACCTTGAACCTATTCATGGTGTCCC 1244
1068 TAATCTATTTTATAGAACACTATCAGACCCCTTCTTCCGAAGATCCGATAATATTATGCC 1127
1245 TACTGTTAGAATTAATTTTAGGAACCTCAGAATACTTTTGAAGAGGTACTGCTAACTA 1304
1128 AACATTAGGAATAATGTAGTCGAGGGGTAGGATTCATTCAACCAATATGGTGAAGT 1187
1305 TAGTCAACCTATAGTCACTGGGCTTCAATTAAGAATTCAGAACTGAATTAACACC 1364
1188 TCTATATAGAGGAGGAAAGTAGATTTCTCTGATGAGTTGCCAATTGA-----CGG 1241
1365 AGAAACACAGAACGACCAATATGAATCATATAGTCATAGGTTATCTCATAAGGGCT 1424
1242 TGAGAAATTCATTAGTTGGATATAGTCATAGATTAAAGTCACTGATTAACCAAGGCTGT 1301
1425 CATTTCAACAATCTAGGGTCATGTACCAGTATATTCTTGGACGACCGTAGTGACAGATCG 1484
1302 ATATTAATTAATATACTAGCTTGGCAACATTTGTTGGACACATCACAGTGTACTGA 1361
1485 TACAATACCAATTTAGTTTCAGATAGCATAAACAAATACCATTTGGTAAATCAATCAACT 1544
1362 TCGAAATATACTATCCGATGTAATTACACAAATACCATTTGGTAAATCAATCTCCCT 1421
1545 TAAATTCAGGTACTCTGTAGTCAGTGGCCAGGATTTACAGAGGGGATTAATCCGAAC 1604
1422 TACTTCAGGTACTCTGTAGTCAGAGGCCAGGATTTACAGAGGGGATATCATCCGAAC 1481
1605 TAACTGTAATGTAGTGTACTAAGTATGGTCTTAATTTTAAATAATCATATCAATACAGCG 1664
1482 TAACTGTAATGTAGTGTACTAAGTATGAGTCTTAATTTTGTATATATCATATACAGCG 1541
1665 GTATCCGCTGAGAGTTCGTTATGCTCTCAAAATGGTCTCTGAGGGTAACTGTCCG 1724
1542 GTATCCGCTGAGAGTTCGTTATGCTCTCAAAATGGTCTATGAGAGTAAATGTTGG 1601
1725 AGGGAGTACTACTTTTGTATCAAGGATTCCTAGTACTATGAGTCAAAATGAGTCTTTGAC 1784
1602 AGGGAGTACTACTTTTGTATCAAGGATTCCTAGTACTATGAGTCAAAATGGTCTTTGAC 1661
1785 ATCTCAATCATTTAGATTGACAGAAATTCCTCTAGTATAGTGCATCTGGCAGTCAAC 1844
1662 ATCTCAATCATTTAGATTGACAGAAATTCCTCTAGGCAATAGTATCTGGCAGTCAAC 1721
1845 TGCTGGAATTAAGTATAAGTAAATGACAGGTAGACAAACGTTTCACTTTGATAAAATGA 1904
1722 TGCTGGAATTAAGTATAAGTAAATCCAGGTAGACAAACGTTTCACTTTAGATAGATTGA 1781
1905 ATTCAATTCGAATTTACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAGAGGC 1964
1782 ATTTATCCAGTGTGATGCAACATTTTGAAGCAGAAATATGATTTAGAAAGAGCACAAGGC 1841
1965 GGTGAATGCTCTGTTTACTTAATACGAATCCAGAGATTGAAACAGAGATGTGACAGATTA 2024
1842 GGTGAATTCGCTGTTTACTTCTTCCAAATCAATCGAGTTTAAACAGATGTGACGATTA 1901
2025 TCATATTGATCAAGTATCCAAATTTAGTGGCGGTTTATCGGATGAATTTCTGCTTAGATGA 2084

1902 TCATATTGATCAAGTATCCAAATTTAGTAGATTGTTATCCGATGAATTTGTCTGATGA 1961
2085 AAAGAGAGAAATTAATTGAGAAAGTGAATATCGGAAACGACTCAGTGATGAAGAAACTTT 2144
1962 AAAGCAGAAATTTGTCGAGAAAGTCAAAATCGGAAAGGACTCAGTGATGAGCGGAATTT 2021
2145 ACTCCAGATCCAAACTTTCACTCATCAATAAGCAACAGACTTTCAATCTCTACTTAATGA 2204
2022 ACTTCAAGATCCAAACTTTCAAGGGATCAATAGGCACAGAC----- 2064
2205 GCAATCGAATTTCACTATCTATCATGAACAATCTGAAACATGGATGGTGGGGAAGTGAGAA 2264
2065 -----CGTGGCTCGAGAGGAAGTACGGA 2087
2265 CATTACATCCAGGAAGAAATGACGTATTTAAAGAGAAATTAACGTCACTACCGGGAC 2324
2088 TATTACCATCAAGAGGAGATGACGTATTCAGAGAGAAATTAACGTCACTACCAAGGTAC 2147
2325 TTTTAATGAGTGTATCCGAGCTATTTATATCAAAAAATAGGAGTCCGGAATTTAAAGC 2384
2148 CTTTGTAGTGTCTATCCAGCTATTTGTATCAAAAAATAGATGAGTCCGAAATTTAAAGC 2207
2385 TTATACTCGCTACCAATTAAGAGGCTATTTTGAAGATAGTCAAGATTTTAGAGATATATTT 2444
2208 CTATAACCGTTTACCAATTTAAGAGGCTATTCGAAGATAGTCAAGACTTTAGAAATCTATTT 2267
2445 GATTGTTATTAATCGGAAACATGAACATTTGGATGTTTCCAGGTACCGAGTCCGTATGGCC 2504
2268 AATTGCTTACAAATGCAAAACACGAACAGTAAATGTACAGGTACCGGTTCTTATGGCC 2327
2505 GCTTTCAGTTCAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGCACCACTTT 2564
2328 GCTTTCAGTTCGAAAGTCCAAATTTGGAAGTGGGAGAACCGAATCGGTGTGCGCACACT 2387
2565 TGAATGGAATCCTGATCTAGATTGTTCTCGAGAGATGAGAGAAATTTGCGGCATCATTC 2624
2388 TGAATGGAATCCTGATTTAGATTGTTCTCGAGAGACGGGGAATAATGTGCACATCATTC 2447
2625 CCATCATTTCTCTTTGATATTTGATTTGATGATGACAGACTTGCATGAGAAATCTAGGGCT 2684
2448 CCATCATTTCTCTTTGATTTGATTTGATGATGATGATGATGATGATGATGATGATGAT 2507
2685 GTGGGTGTATTTCAAGATTTAAGACGAGGAGGTCATGCAAGCTAGGGAATCTGGAATTT 2744
2508 GTGGGTGTATTTCAAGATTTAAGACGAGGAGGTTATGCAAGATTTAGGAATTTCTGGAATTT 2567
2745 TATTGAAGAGAAACCATTTATTAGGAGAGCACTGTCTGCTGTGAAGAGAGAGAGAGAGAAA 2804
2568 TATCGAAGAGAAACCATTTAATTTGGAGAGCACTGTCTGCTGTGAAGAGAGAGAGAGAAA 2627
2805 ATGAGAGACAAACGTTGAATACTCAATTTGGAAACAAACAGAGTATATACAGAGGCAAA 2864
2628 ATGAGAGACAAACGTTGAATACTCAATTTGGAAACAAACAGAGTATATACAGAGGCAAA 2687
2865 AGAAGCTGTGATCTTTTATTTTGTAGATTTCTCAATATAATAGATTTACAAGCGGATACAAA 2924
2868 AGAAGCTGTGATCTTTTATTTTGTAGATTTCTCAATATAATAGATTTACAAGCGATACAAA 2747
2925 CATTGGCATGATTCAATCGGAGATAAATCTTGTTCATCGAATTCGAGAGGCTTATCTGTC 2984
2748 CATTGGTATGATTCAATCGGAGATAGACTTGTTCATCAGATCCACAGGCTTATCTTCC 2807
2985 AGAATTAATCTGTTATCCGGGTGTAATGCGGAAATTTTGAAGAAATTTAGAGGTCGAT 3044
2808 AGAATTAATCTTCAATCCAGGAAATAATGTGGTGTGATTTTGAAGAAATTTAGAAACCGGAT 2867
3045 TATCACTGCAATCTCCCTATACGATCGAGAAATGTCTGTTAAATGTGTTAAATTTTAAATA 3104
2868 TTCTACTGCAATTAATCCCTATATGATGAGGAAATGTCAATTAATAATGCGGATTTCAATTA 2927
3105 TGGATTAGCATGCTGGAATGTAAAGGCGATGTAGAT---GTACAACAGAGCCATCACCG 3161

```
Db 2928 TGGCTTATCATGCTGGAACTGAAAGGCGATGTAGATGTAGTAGAACAACCAACACCG 2987
Qy 3162 TTCTGCTCTTGTATCCAGAAATGGGAGCAGAGTGTCAACAGCAGTTCGCGTCTGTCC 3221
Db 2988 TTCTGCTCTTGTATCCAGAAATGGGAGCAGAGTGTCAACAACTTGTGTCTGTCC 3047
Qy 3222 GGGCGTGGCTATATCTCTCGTGTCAAGCGGTACAAAGAGGGATATGGAGAGGTTGTGT 3281
Db 3048 GGGCGTGGCTATATCTCTCGTGTACAGCGTACAAAGAGGGATATGGAGAGGTTGTGT 3107
Qy 3282 AAGATCCATGAATCAGAACCAATACAGACGAACCTAAATTTTAAAACTGTGAAGAAGA 3341
Db 3108 AAGATCCATGAATCAGAACCAATACAGACGAACCTAAATTTTAAAACTGTGAAGAAGA 3167
Qy 3342 GGAAGTGTATCCACGATACAGGACGTTGATGATGATGATGATGATGATGATGATGATG 3401
Db 3168 GGAAGTGTATCCACGATACAGGACGTTGATGATGATGATGATGATGATGATGATGATG 3227
Qy 3402 AG-----TATGTAATTCCTGTAATGCTGGATATGAGGATGATGAGTGCATATGAAGTTGA 3449
Db 3228 AGATCCACAGATTCATGTAATTCCTGTAATGATATGATGATGATGATGATGATGATGATG 3287
Qy 3450 TACTACAGATCTGTTAAATACAAACCGACTTATGAAGAAGAAACGTTATACAGATGTACG 3509
Db 3288 TACTACAGATCTGTTAAATACAAACCGACTTATGAAGAAGAAAGGTATACAGATGTACA 3347
Qy 3510 AAGGATTAATCTGTGAATATGACAGAGGTTATGTAATATCCACCACTACAGCTGG 3569
Db 3348 AGGAGATAATCTGTGAATATGACAGAGGTTATGTAATATCCAGCTACAGCTGG 3407
Qy 3570 TTATATGACAAAGAAATAGAACTATCCAGAAACCGATAAGGTATGAGATTGAGATTGG 3629
Db 3408 TTATGTAAGAAAGAAATAGAACTATCCAGAAACCGATAAGGTATGAGATTGAGATTGG 3467
Qy 3630 AGAAACGGAAGGAAATTTATTTAGACAGCGTGAATTTACTCTTATGGAGGAA 3684
Db 3468 AGAAACGGAAGGAAATTTATTTAGACAGCGTGAATTTACTCTTATGGAGGAA 3522
```

RESULT 14

```
PCT-US95-10310-1
; Sequence 1, Application PC/TUS9510310
; GENERAL INFORMATION:
; APPLICANT: MYCOGEN CORPORATION 5501 Oberlin Drive
; APPLICANT: STREET ADDRESS: 5501 Oberlin Drive
; APPLICANT: CITY: San Diego
; APPLICANT: STATE/PROVINCE: California
; APPLICANT: COUNTRY: US
; APPLICANT: POSTAL CODE/ZIP: 92121
; APPLICANT: PHONE NUMBER: (619) 453-8030
; APPLICANT: FAX NUMBER: (619) 453-6991
; TITLE OF INVENTION: Protein Toxins Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10310
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; CLASSIFICATION:
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MAS0.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: LambdaGem - 11 (tm) Library of August Sick
; CLONE: 81A2
; PCT-US95-10310-1
```

```
Query Match 48.6%; Score 1793.4; DB 5; Length 3522;
Best Local Similarity 71.9%; Pred. No. 0;
Matches 2497; Conservative 0; Mismatches 861; Indels 117; Gaps 7;

Qy 228 ACAATAGCTAGCTTTTATATAGTTTCTTGTGTGTAATTTATGGCCCGCGCAGAGATCA 287
Db 147 ACTAGGGGATTTATATCTTGGCTTGTTCATGTAATATGGGGGCTATAGGTCCTTCACA 206
Qy 288 GTGGGAAATTTCTTAGAACATGTGCAACAACCTTATAAATCAACAATTAACAGAAATGC 347
Db 207 ATGGGATATATTTTATAGACAAATTTAGCTATTTGATCGGCCAAAGAAATAGAGAAATTCGC 266
Qy 348 TAGGAATACGGCACTTGTCTCGATTACAAGTTTATAGGAGATTCCTTTAGAGCCTTATCAACA 407
Db 267 TAGGAATCAGGCATTTCTAGATTACAGGGCTAGCAATCTTTACCGAATTTACACAA 326
Qy 408 GTCACCTTGAAGATTTGGCTAGAAACCGTGATGATGCAAGAACGAGAGAGTGTCTTTTATAC 467
Db 327 TGCTTTTAAAAAATGGGAAAGTAGATCTCTACTAATCCAGCATTAAGAGAGAGAGATGCGTAT 386
Qy 468 CCAATATATAGCTTAGAACCTTGAATTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAA 527
Db 387 TCAATTTAATGACATGAACAGTCTCTTACAACAGCTATTTCTCTTTTTCAGTTCAAGG 446
Qy 528 CCAAGAAAGTTCCATTTAATAGTATATGCTCAAGCTGCAAAATTTTACACCTATTTATTT 587
Db 447 TTATGAAATTCCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTTACATTTATCGGTTT 506
Qy 588 GAGAGATGCTCTCTTTTGTAGTGAATTTGGGCTTACATCGCAGAAATTCACAGTTA 647
Db 507 GAGAGATGTTTCAGTGTTTGGCAACAGTTGGGATTTGATGTAGCAACAATCAATAGTCG 566
Qy 648 TTATGAGCCCAAGTGGAAACAAACAGAGATTTATCCGACTATTGCGTAGAATGGTATAA 707
Db 567 TTATAATGATTTAACTAGGCTTATTTGGCGAATATACTGATTTATGCTGTGTTGGTATAA 626
Qy 708 TACAGGTCTAAATAGCTTCGAGAGGGACAAATGCGC--CAAGTTGGTGGCTTATATAATCA 764
Db 627 TACGGGTTAAATCGTTTACCAGTAAATGAAGGGTACGAGGATGGCAAGATTTAATAG 686
Qy 765 ATTCCGTAGAGATCTAACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGA 824
Db 687 GTTTAGAGAGAGGTTAAACAATATCAGTATTAGATATTTATTTCTTTTTCAAAATACGA 746
Qy 825 CACTCGCACTTATCCAAATTAATACAGAGTGTCTAGTTAAACAGGGAAGTTTATACAGCGC 884
Db 747 TTCTAGATTATATCCAAATTCGGAATTCGGAATCTATCAATTAACCGGGAAGTATATACAGATCC 806
```

QY 885 AATTGGAGCAACAGGGGTAATAATATGCAAGTATGAAATGGTATATAATAATAATATGCACTTC 944
DB 807 -----GTAATTAATAATACTGATTAATAGATTACCCCAAG 842
QY 945 GTTTTCGGCTATAGAGCTGGGTTATCCGAGCCGGCATCTACTTGTATTTCTAGAAC 1004
DB 843 TTTTCAGAGTATGAAATTCAGCTATTAAGAGTCCCATCTTATGGGATTTCTTAATAA 902
QY 1005 ACTTACAAATTTTAGCACCTTCATCAGATGGAGTGTCTACTAGGCATATGACTTACTGGCG 1064
DB 903 TATAATTTGACACTGATTA-----ATTAGGGGTTCACTATTGGGC 947
QY 1065 GGGGACACAAATCAATCTCGGCCAATAGGAGCGGATTAATACTCAACGATGGGTC 1124
DB 948 GGGGCATCGTGTAACTTCTCAATTTACCGGTAGTTTCGCAAGTGAAGCTCCCTCAATA 1007
QY 1125 TACCAATACTTCTATTAATCTGTGAAGATTATCATCTTCTCTCGAGAGTATATTGGAC 1184
DB 1008 CGGGATAACTGCANAACGCAAGACCGAGTGAATATGCTCTGACACTTTTCCAGGTCT 1067
QY 1185 TGAATCATATGACGAGTGTCTTATGGGAATTTTACCTTGAACCTTATCATGGTGTCC 1244
DB 1068 TAATCTATTTATAGAACACTATCAGACCTTCTTCCGAGATCGATATATTATGCC 1127
QY 1245 TACTGTATAGATTAATTTTAGAACCTTCAGAATACCTTTTGAAGAGGTACTGTAACTA 1304
DB 1128 AACATTAGGAATAATGTAGTGCAGGGGTAGGATTCATTTCAACCAATAATGGTGAAGT 1187
QY 1305 TAGTCAACCTATAGTACCTGGGCTTCAATTAAGATTCAGAACTGAATTACCACC 1364
DB 1188 TCTATATAGAAGGAGGAACAGTAGATTTCTTCTGATGAGTTGCCAATTTGA-----CGG 1241
QY 1365 AGAAACAACAGAACGACCAAAATTAATGAATCATATAGTCATAGTTTATCTCATAGGGCT 1424
DB 1242 TGAGAAATCATTTAGTTGGATATAGTATAGTATAGTCACTTACGTTAAACAGGTCTGT 1301
QY 1425 CATTTCACAATCTAGGGTGCAATACCAAGTATATTTTGGAGCCCGTAGTGACAGATCG 1484
DB 1302 ATATAATACTAATAATACTAGCTTGCCAAACATTTGTTGGACACATCACAGTGTACTGA 1361
QY 1485 TACAATACCATTTAGTTCAGATAGATTAACACAAATACCATTTGGTAAATCATCAACT 1544
DB 1362 TCGAAATATAATCTATCCGGATGTAATTAACAAATACCATTTGGTAAATCATTTCCCT 1421
QY 1545 TAAATTCAGTACTCTGTAGTCAGTGGCCAGGATTTACAGAGGGGATATAATCCGAAC 1604
DB 1422 TACTTCAGTACTCTGTAGTCAGAGGCCAGGATTTACAGAGGGGATATCATCCGAAC 1481
QY 1605 TAACTGTTAATGGTGTACTAAGTATGGGCTTAAATTTTAATAATACATCATTTACAGCG 1664
DB 1482 TAACTGTTAATGGTGTACTAAGTATGAGTCTTAATTTTGAATATACATCATTTACAGCG 1541
QY 1665 GTATCGGTGAGAGTTGTTATGTCTCTCAAAATGTTCTGAGGGTAACTGTCCGG 1724
DB 1542 GTATCGGTGAGAGTTGTTATGTCTCTCAAAATGTTCTGAGAGGATTAATGTTGG 1601
QY 1725 AGGGAGTACTCTTTTGTATCAAGGATTCCTGTAGTACTATGAGTCAAAATGAGTCTTTGAC 1784
DB 1602 AGGGAGTACTCTTTTGTATCAAGGATTCCTGTAGTACTATGAGTCAAAATGAGTCTTTGAC 1661
QY 1785 ATCTCAATCATTTAGATTGCAAAATTTCTGTAGGTTATGATGATCTGGCAGTCAAAC 1844
DB 1662 ATCTCAATCATTTAGATTGCAAAATTTCTGTAGGATTTCTGTAGGATTTAGTATCTGGCAGTCAAAC 1721
QY 1845 TGCTGGAATAGTATAAGTATAATGAGGTAGACAAACGTTTCACTTTGATAAATTTGA 1904
DB 1722 TGCTGGAATAGTATAAGTATAATGAGGTAGACAAACGTTTCACTTTAGTATAAATTTGA 1781
QY 1905 ATTTCATCCAAATCTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAGAGGC 1964
DB 1782 ATTTATCCAGTTTGTATGCAACATTTGAAGCAGAAATATGATTTAGAAAGCAGACAAAGGC 1841
QY 1965 GGTGAATGCTCTGTTTACTTAATACGAATCCGAAGATTTGAACACAGATGTGACAGATTA 2024

DB 1842 GGTGAATTCGCTGTTTACTTCTTCAATCAAAATCGAGTTAAAAACAGATGTGACGGATTA 1901
QY 2025 TCATATTTGATCAAGTATCCAAATTTAGTGGCGTGTATTCGGATGATTTCTGCTTAGATGA 2084
DB 1902 TCATATTTGATCAAGTATCCAAATTTAGTAGATTGTTTATCCGATGAATTTTGTCTGATGA 1961
QY 2085 AAAGAGAGAAATTTACTTTGAGAAAGTGAATAATGCGAAACGACTCAGTGTATGAAGAAACTT 2144
DB 1962 AAAGGAGAAATTTGTCGAGAAAGTCAACATGCGAAGGACTCAGTGTATGAGCGGAATTT 2021
QY 2145 ACTCAAGATCCAAATTTCAATCCATCAATTAAGCAACACGACTTTCATATCTTACTTAATGA 2204
DB 2022 ACTTCAAGATCCAAATTTCAAGAGGATCAATAGGCAACACGAC----- 2064
QY 2205 GCAATCGAATTTCACTATCTATCCATGAACAATCTGAACATGATGGTGGGGAAGTGAGAA 2264
DB 2065 -----CGTGGCTGGAGAGGAAGTACGGA 2087
QY 2265 CATTACAAATCCAGGAAGGAATGACGTATTTAAAGAGAAATTCAGTCAACATCCGCGGAC 2324
DB 2088 TATTTACCAATCCAGAGAGGAGATGAGTATTTCAAGAGAAATTTACGTCACACTTACCAGGTAC 2147
QY 2325 TTTTAATGAGGTATTCGACGTATTTATATATCAAAAAATAGGAGAGTCCGAAATTTAAAGC 2384
DB 2148 CTTTGTAGTGTCTATCCACGTATTTGTATCAAAAAATAGATGAGTCCGAAATTTAAAGC 2207
QY 2385 TTATCTCGCTACCAATTTAGAGGGTATTTGAAGATAGTCAAGATTTAGAGATATATTT 2444
DB 2208 CTATAACCGTTTACCAATTTAGAGGGTATTCGAAGATAGTCAAGACTTTAGAAATCTATTT 2267
QY 2445 GATTCGTTTATATGCGAAACATGAAACATTCGATGTTTCCAGGTACCGAGTCCGATGACC 2504
DB 2268 AATTCGCTACATTCGAAAAACAGAAACAGTAATTTGACAGGTACCGGTTCTTTATGGCC 2327
QY 2505 GCTTTCAGTGTAAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCAACCACTTT 2564
DB 2328 GCTTTCAGTGTAAAGTCCAAATTCGAAGGTGCGGAGAACCGAATCGGTGTGTGCCACACT 2387
QY 2565 TGAATGAAATCCTGATCTAGATTTGTTCTGCGAGAGATGGAGAAAAATGCGGCATCATTC 2624
DB 2388 TGAATGAAATCCTGATTTAGATTTGTTCTGCGAGAGCGGGGAAAAATGTCACACTCATTC 2447
QY 2625 CCATCATTTCTTTGGATATTTGATATTTGATGACACAGACTTGCATGAGAACTTAGGGCT 2684
DB 2448 CCATCATTTCTTTGGACATTTGATTTGGATGACACAGACTTGCAGAGGATCTTAGGGCT 2507
QY 2685 GTGGGTGATTTCAAGATTTAAGACGCGGAAAGGTCTCAAGACTTAGGAAATCTGGAATTT 2744
DB 2508 GTGGGTGATTTCAAGATTTAAGACGCGGAAAGGTATGCAAGATTTAGGAAATCTGGAATTT 2567
QY 2745 TATTGAAGAGAAACCATTTATAGGAGAAAGCTGTCTGCTGTGAAGAGAGAGAGAGAAAA 2804
DB 2568 TATCGAAGAGAAACCATTTAATTTGAGAAAGCACTGTCTCTGTGTGAAGAGCGGAAAAAAA 2627
QY 2805 ATGAGAGACAAACGTCGAAAACTTACAAATTCGAAAAACAGAGTATATACAGAGGCAAA 2864
DB 2628 ATGAGAGACAAACGCGGAAAAAATCAATTTGAAAAACAAACAGATATATACAGAGGCAAA 2687
QY 2865 AGAAGCTGTGGATGCTTTATTTGTAGATTTCAATATAATAGATTTACAAGCGGATACAAA 2924
DB 2688 AGAAGCTGTGGATGCTTTATTTCTGTAGATTTCAATATGATAGATTTACAAGCAGATACAAA 2747
QY 2925 CATTTGGCATGATTCATCGCGCAGATTAACCTTGTTCATCGAATTCGAGAGGCTTATCTGTC 2984
DB 2748 CATTTGGTATGATTCATCGCGCAGATAGACTTTGTTTCATCAGATCCAGAGGCTTATCTTCC 2807
QY 2985 AGAATTAATCTGTTATCCGGGTGTAATTCGCGAAATTTTGAAGAAATTTAGAAAGTCCGAT 3044
DB 2808 AGAATTAATCTGTTATCCGGAATTAATGTTGGTGTGATTTTGAAGAAATTTAGAAACCGTAT 2867
QY 3045 TATCACTGCAATCTCCCTTATACGATGCGAGAAATGTCGTTAAAAATGGTGTATTTAATAA 3104

Db 1302 ATATAATACTAATAATACTAGCTTGCCAACTTTGTTGGACACATCAAGTCTACTGA 1361
Qy 1485 TACAATAACCAATAGTTACAGATAGACATAACACAAATACCAATGTTAAATCAATCAACT 1544
Db 1362 TCGAATAATAACTATCCGATGTAAATACACAAATACCAATGTTAAATCAATCTCCCT 1421
Qy 1545 TAAATTCAGGTACTCTCTGTAGTCAAGTGGCCAGGATTTACAGAGGGGATATAATCCGAAC 1604
Db 1422 TACTTCAGGTACTCTCTGTAGTCAAGGCGCCAGGATTTACAGAGGGGATATCATCCGAAC 1481
Qy 1605 TAACTGAATGAGTGTACTAGTATAGTGTCTTAATTTTAATAATACATCATACAGG 1664
Db 1482 TAACTGAATGAGTGTACTAGTATAGTGTCTTAATTTTAATAATACATCATACAGG 1541
Qy 1665 GTATCGCGTGAAGTTCGTTATGCTCTCTCAAAACAATGGTCTGAGGTTAACTGTGCG 1724
Db 1542 GTATCGCGTGAAGTTCGTTATGCTCTCTCAAAACAATGGTCTGAGTGAAGTGTGG 1601
Qy 1725 AGGAGTACTACTTTTGATCAAGGATCCCTAGTACTATGAGTGCATAATGAGTCTTTGAC 1784
Db 1602 AGGAGTACTACTTTTGATCAAGGATCCCTAGTACTATGAGTGCATAATGAGTCTTTGAC 1661
Qy 1785 ATCTCAATCAATTTAGATTTGCAAGATTTCTGCTAGTATAGTGCATCTGGCAGTCAAC 1844
Db 1622 ATCTCAATCAATTTAGATTTGCAAGATTTCTGCTAGTATAGTGCATCTGGCAGTCAAC 1721
Qy 1845 TGCTGGAATAAGTATAAGTAATAATGCAAGTACAGAAAGTTCACCTTTGATAAATTTGA 1904
Db 1722 TGCTGGAATAAGTATAAGTAATAATGCAAGTACAGAAAGTTCACCTTTGATAAATTTGA 1781
Qy 1905 ATTCAATCCAAATTTACTGCAACCTTCGAAGCAGAATACGATTTAGAAAGGCGCAAGAGC 1964
Db 1782 ATTTATCCAGTTGATGCAACATTTGGAAGCAGAAATGATGATTTAGAAAGCAGCAAAAGC 1841
Qy 1965 GGTGAATGCTCTGTTTACTTAATACCAATCCAAAGAGATTTGAAAACAGATGTGACAGATTA 2024
Db 1842 GGTGAATGCTCTGTTTACTCTTCCAAATCCAAATCGAGTTTAAAAACAGATGTGACAGATTA 1901
Qy 2025 TCATATTCATCAAGTATCAATTTAGTGGCGTTTATCGGATGAATTTCTGTTAGATGA 2084
Db 1902 TCATATTCATCAAGTATCAATTTAGTGGCGTTTATCGGATGAATTTCTGTTAGATGA 1961
Qy 2085 AAAGAGAGAAATTTGGAAGAGTGAATATGCGAAACAGCTCAGTGTAGTGAAGAACTT 2144
Db 1962 AAAGCAGAAATTTGCGGAAGAGTCAACATGCGAAGCGACTCAGTGTAGTGGCGAAATTT 2021
Qy 2145 ACTCCAAGTCCAACTTCATCCATCAATGAACCAACAGACTTCATATCTAATGA 2204
Db 2022 ACTCCAAGTCCAACTTCAGAGGATCAATAGGCAACAGAC----- 2064
Qy 2205 GCAATCGAATTTACATCTATCCATGAACAATCTGAACATGATGCTGGGAAGTGAGAA 2264
Db 2065 -----CGTGGCTGGAGAGGAGTACGGA 2087
Qy 2265 CATTAACAATCCAGGAAGGAATGAGTATTTAAAGAGAAATACGTCACTACCGGGAC 2324
Db 2088 TATTACCATCCAAGGAGGATGACGATTTCAAAGAGAAATACGTCACTACCAAGGTAC 2147
Qy 2325 TTTTAATGAGTGTATCCGACGATTTATATCAAAAAATAGAGAGTCCGAAATTTAAAGC 2384
Db 2148 CTTTGATGAGTGTATCCAAACGATTTGATCAAAAAATAGATGAGTCCGAAATTTAAAGC 2207
Qy 2385 TTATACTCGTACCAATTAAGAGGATATTTGAAGATAGTCAAGATTTAGAGATATATTT 2444
Db 2208 CTATAACCGTTTACCAATTAAGAGGATATCGAGATAGTCAAGACTTTAGAAATCTATTT 2267
Qy 2445 GATTCGTTATATGCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGGCC 2504
Db 2268 AATTCGCTACAAATGCAAAACAGAAACAGTAAATGTCACAGGTACCGGTTCCTTATGGC 2327
Qy 2505 GCTTTGAGTGAAGCCCAATCGGAAGTGGGAGAACCGAATCGATGGCAGCACTTT 2564
Db 2328 GCTTTGAGTGAAGTTCGAATTTGGAAGTGGGAGAACCGAATCGATGGTGTGTCAGACCT 2387

Qy 2565 TGAATGGAATCCTGATCTAGATTTGTTCTCGAGAGATCGAGAAAATGTGCGCATCTTC 2624
Db 2388 TGAATGGAATCCTGATTTAGATTTGTTCTCGAGAGATCGGGAANAATGTGCATCATCTC 2447
Qy 2625 CCATCATTTCTCTTTGGATATTTGATATTTGGATGCAAGACTTTGCATGAGAATCTAGGCGT 2684
Db 2448 CCATCATTTCTCTTTGGACATTTGATTTGGATGCAAGACTTTGCAAGAGGATCTAGGCGT 2507
Qy 2685 GTGGGTGATTTCAAGATTTAAGACGCAAGAGGTCTCAAGACTTAGGGAATCTGGAATTT 2744
Db 2508 GTGGGTGATTTCAAGATTTAAGACGCAAGAGGTCTCAAGATTTAGGAATCTGGAATTT 2567
Qy 2745 TATTGAAGAGAAACCAATTTATTAGGAGAGCACTGTCTCGTGAAGAGACGAGAGAAA 2804
Db 2568 TATCGAGAGAAACCAATTTATTGGAGAGCACTGTCTCGTGAAGAGCGGGAANA 2627
Qy 2805 ATGAGAGACAAACGTTGAAAACTACAATTCGAAAACAAAACGAGTATATACAGAGCAAA 2864
Db 2628 ATGAGAGACAAACGCGGAAAACTACAATTCGAAAACAAAACGAGTATATACAGAGCAAA 2687
Qy 2865 AGAAGCTGTGATGCTTTTATTGTTAGATTTCTCAATATATAGATTTCAAGCGGATACAAA 2924
Db 2688 AGAAGCTGTGATGCTTTTATTGTTAGATTTCTCAATATATAGATTTCAAGCGGATACAAA 2747
Qy 2925 CATTTGGCATGATTCATGCGGACAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTGTC 2984
Db 2748 CATTTGGTATGATTCATGCGGACAGATAGACTTTGTTTCATCAGATCCAGAGGCTTATCTTCC 2807
Qy 2985 AGAATATCTGTTATCCGGGTGTAATTCGCGAATTTTGAAGAAATTTAGAGGTCGAT 3044
Db 2808 AGAATACCTTTTCATTTCCAGGAATAAATGTGCTGATTTTGAAGAAATTTAGAAAACCGTAT 2867
Qy 3045 TATCACTGCAATCTCCCTATACGATCGGAGAAATGTCGTTAAAAATGCTGATTTTAAATA 3104
Db 2868 TTCTACTGCATTTATCCCTATATGATGCGAGAAATGTCATTTAAAAATGCGGATTTCAATA 2927
Qy 3105 TGGATTAGCATGCTGGGAATGTAAAAGGCACTGTAGAT---GTACAACAGAGCCATCACCG 3161
Db 2928 TGGCTTATCATCTGGAACGTTGAAAGGCACTGTAGATGTAGTAGNACAAAACACACCG 2987
Qy 3162 TTCTGCTCTGTTATCCAGAAATGGGAAGCAGAGTGTCAAGCAGTTCGCGTCTGTC 3221
Db 2988 TTCTGCTCTGTTATCCCGGAAATGGGAAGCAGAGTGTCAAAAATAATTCGTTCTGTC 3047
Qy 3222 GGGGCGTGGCTATATCTCCGTTGTCAGCGCTACAAAGAGGATATGAGAGGTTGTTGT 3281
Db 3048 GGGGCGTGGCTATATCTCCGTTGTTACAGCGCTACAAAGAGGATATGAGAGGTTGCGT 3107
Qy 3282 AACGATCCATGAAATCGAGAAACAATACAGACGAACTAAAAATTTAAAAACTGTGAAGAAGA 3341
Db 3108 AACCATCCATGAGATCGAGAAACAATACAGACGAACTAAAAATTTAAAAACTGTGAAGAAGA 3167
Qy 3342 GGAAGTGTATCCAAACGATACAGAAACGTTGATTAATGATTTATCTGCACACCAAGGTACAGC 3401
Db 3168 GGAAGTGTATCCAAACGATACAGAAACGTTGATTAATGATTTATCTGCACACCAAGGTACAGC 3227
Qy 3402 AG-----TATGTAATTCGCTGATGCTGGAATGATGAGGATGATGAGTTGA 3449
Db 3228 AGGATCCACAGATTCATGTAAATTCCTGTAATATCAGATATGAGGATGATGAAATGAA 3287
Qy 3450 TACTACAGCATCTGTTAAATTAACAAACGACCTTATGAAGAAGAAACGATATACAGATGTAAG 3509
Db 3288 TACTACAGCATCTGTTAAATTAACAAACGACCTTATGAAGAAGAAAGGATATACAGATGTAAG 3347
Qy 3510 AAGAGATAATCATTTGTAATATGACAGAGGATATGTAATTTATCCACATACCAAGCTGG 3569
Db 3348 AGGAGATAATCATTTGTAATATGACAGAGGATATGTAATTTATCGACAGTACCAAGCTGG 3407
Qy 3570 TTATATGACAAAGAAATTTAGNATACCTCCAGAAACCGATAGGATGATGAGATTGG 3629
Db 3408 TTATGTGACAAAGAAATTTAGAGTACTCTCCAGAAACCGATAGGATGATGAGATTGGATCGG 3467

us-10-614-524-1.rni

Qy 3630 AGAAACGGAGGAAAGTTTATTCTAGACACGCTGGAATTACTCCTTATGGAGGAA 3684

Db 3468 AGAAACGGAGGGAAGTTTATTCTAGACAAATGTCGAATTACTCCTTATGGAGGAA 3522

Search completed: November 21, 2004, 16:55:54
Job time : 355 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 08:55:06 ; Search time 1733 Seconds
(without alignments)
11496.341 Million cell updates/sec

Title: US-10-614-524-1
Perfect score: 3687
Sequence: 1 ttgacttcaaataggaaaa.....tactcttatggaggaatag 3687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues
Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3687	100.0	3687	16	US-10-614-524-1
2	3666.2	99.4	4173	15	US-10-428-961-37
3	3260.2	88.4	3684	15	US-10-428-961-62
4	3133.6	85.0	3687	17	US-10-809-953-9
5	1581.6	42.9	3558	9	US-09-826-660-22
6	1495.6	40.6	3624	10	US-09-988-462-6
7	1408.4	38.2	3522	9	US-09-826-660-5
8	1408.4	38.2	3522	11	US-09-837-961-7
9	1408.4	38.2	3522	18	US-10-825-751-7
10	1402	38.0	3534	9	US-09-873-873-25
11	1402	38.0	3534	10	US-09-916-956A-25
12	1402	38.0	3534	10	US-09-997-914-25

13	1402	38.0	3534	15	US-10-365-645-25	Sequence 25, Appl
14	1402	38.0	3534	16	US-10-672-163-25	Sequence 25, Appl
15	1402	38.0	3534	17	US-10-739-482-25	Sequence 25, Appl
16	1402	38.0	3534	17	US-10-817-182-25	Sequence 25, Appl
17	1400.6	38.0	3531	9	US-09-873-873-9	Sequence 9, Appl
18	1400.6	38.0	3531	9	US-09-873-873-11	Sequence 11, Appl
19	1400.6	38.0	3531	9	US-09-873-873-13	Sequence 13, Appl
20	1400.6	38.0	3531	10	US-09-916-956A-9	Sequence 9, Appl
21	1400.6	38.0	3531	10	US-09-916-956A-11	Sequence 11, Appl
22	1400.6	38.0	3531	10	US-09-916-956A-13	Sequence 13, Appl
23	1400.6	38.0	3531	10	US-09-997-914-9	Sequence 9, Appl
24	1400.6	38.0	3531	10	US-09-997-914-11	Sequence 11, Appl
25	1400.6	38.0	3531	10	US-09-997-914-13	Sequence 13, Appl
26	1400.6	38.0	3531	15	US-10-365-645-9	Sequence 9, Appl
27	1400.6	38.0	3531	15	US-10-365-645-11	Sequence 11, Appl
28	1400.6	38.0	3531	15	US-10-365-645-13	Sequence 13, Appl
29	1400.6	38.0	3531	16	US-10-672-163-9	Sequence 9, Appl
30	1400.6	38.0	3531	16	US-10-672-163-11	Sequence 11, Appl
31	1400.6	38.0	3531	16	US-10-672-163-13	Sequence 13, Appl
32	1400.6	38.0	3531	17	US-10-739-482-9	Sequence 9, Appl
33	1400.6	38.0	3531	17	US-10-739-482-11	Sequence 11, Appl
34	1400.6	38.0	3531	17	US-10-739-482-13	Sequence 13, Appl
35	1400.6	38.0	3531	17	US-10-817-182-9	Sequence 9, Appl
36	1400.6	38.0	3531	17	US-10-817-182-11	Sequence 11, Appl
37	1400.6	38.0	3531	17	US-10-817-182-13	Sequence 13, Appl
38	1392.4	37.8	3534	9	US-09-873-873-27	Sequence 27, Appl
39	1392.4	37.8	3534	10	US-09-916-956A-27	Sequence 27, Appl
40	1392.4	37.8	3534	10	US-09-997-914-27	Sequence 27, Appl
41	1392.4	37.8	3534	15	US-10-365-645-27	Sequence 27, Appl
42	1392.4	37.8	3534	16	US-10-672-163-27	Sequence 27, Appl
43	1392.4	37.8	3534	17	US-10-739-482-27	Sequence 27, Appl
44	1392.4	37.8	3534	17	US-10-817-182-27	Sequence 27, Appl
45	1390	37.7	3567	10	US-09-972-175-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-10-614-524-1
; Sequence 1, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boets, Annemie
; APPLICANT: Dame, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vanneste, Stijn
; APPLICANT: Van Rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from Bacillus thuringiensis.
; FILE REFERENCE: NEWBTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3687
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3687)
US-10-614-524-1

Query Match 100.0%; Score 3687; DB 16; Length 3687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGACTTCAAATAGGAAAAATGAGAAATGAAATATATAATGCTTTATCGATCCAGCTGTA 60

1 TCGACTTCAATAGGAAATGAGATGAAATTAATAATGCTTATCGATCCAGCTGTA 60
61 TCGAATCATTTCCACACAAATGGATCTATCAACAGATGCTCGTATGAGGATCTTTGTGT 120
61 TCGAATCATTTCCACACAAATGGATCTATCAACAGATGCTCGTATGAGGATCTTTGTGT 120
121 ATAGCCGAGGGGAATTAATCAATCACTTGTAGGCGATCAACAGTCCAAACGGGTATT 180
121 ATAGCCGAGGGGAATTAATCAATCACTTGTAGGCGATCAACAGTCCAAACGGGTATT 180
181 AACATAGCTGGTAGAATCTAGGTGTATTAGGGGTATCCGTTTCTCGACAAATAGCTAGT 240
181 AACATAGCTGGTAGAATCTAGGTGTATTAGGGGTATCCGTTTCTCGACAAATAGCTAGT 240
241 TTTTATAGTTTTCTTGTGTGTAATTAATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC 300
241 TTTTATAGTTTTCTTGTGTGTAATTAATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC 300
301 CTAGAACATGTCGAACAACTTATTAATCAACAAATACAGAAATGCTAGGATACGGCA 360
301 CTAGAACATGTCGAACAACTTATTAATCAACAAATACAGAAATGCTAGGAAATACGGCA 360
361 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTTATCAACAGTCACTTGAAGAT 420
361 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTTATCAACAGTCACTTGAAGAT 420
421 TGCTAGAAAACCGTGATGATCAAGAAACGAGAAAGTGTCTTTATACCCAAATATATAGCC 480
421 TGCTAGAAAACCGTGATGATCAAGAAACGAGAAAGTGTCTTTATACCCAAATATATAGCC 480
481 TTAGAACTTGAATTTCTTAATGCGATGCGCCTTTTTCGCAATTAGAAACCAAGAAAGTTCCA 540
481 TTAGAACTTGAATTTCTTAATGCGATGCGCCTTTTTCGCAATTAGAAACCAAGAAAGTTCCA 540
541 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTATATTAGAGATGCTCTCT 600
541 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTATATTAGAGATGCTCTCT 600
601 CTTTTTGGTATGAAATTTGGGCTTACATCGCAGAAATTCACAGTTATTATAGAGCGCCAA 660
601 CTTTTTGGTATGAAATTTGGGCTTACATCGCAGAAATTCACAGTTATTATAGAGCGCCAA 660
661 GTGGAAACAAACGAGAGATTAATCCGACTATTTGGTGAATGGTATTAATACAGTCTAAAT 720
661 GTGGAAACAAACGAGAGATTAATCCGACTATTTGGTGAATGGTATTAATACAGTCTAAAT 720
721 AGCTTGAGAGGGACAAATGCGCAAGTTGGGTGCGTTTATTAATCAATTTCCGTAGAGATCTA 780
721 AGCTTGAGAGGGACAAATGCGCAAGTTGGGTGCGTTTATTAATCAATTTCCGTAGAGATCTA 780
781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 840
781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 840
841 ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
841 ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
901 GTAAATATGCGAGTATGAATTTGGTATTAATTAATTAATGACCTTCGTTTTCCGCTATAGAG 960
901 GTAAATATGCGAGTATGAATTTGGTATTAATTAATTAATGACCTTCGTTTTCCGCTATAGAG 960
961 ACTGCGGTTATCCGAAGCCGCGATCTACTTGATTTTCTAGAACAACTTACAAATTTTAGC 1020
961 ACTGCGGTTATCCGAAGCCGCGATCTACTTGATTTTCTAGAACAACTTACAAATTTTAGC 1020
1021 ACTTCATCAGATGGAGTGTCTAGGCAATATGACTTACTGGGGGGGCGACAAATTCAA 1080
1021 ACTTCATCAGATGGAGTGTCTAGGCAATATGACTTACTGGGGGGGCGACAAATTCAA 1080
1081 TCTCGGCCAAATAGGAGCGGATTAATTAATCTCAACGATGGGTCTACCAATATCTTCTATT 1140

1081 TCTCGGCCAAATAGGAGCGGATTAATTAATCTCAACGATGGGTCTACCAATACTTCTATT 1140
1141 AATCTGTGAAGATTAATCATTTCTCTCGAGACGTATATTGCACTGATCATATGACGA 1200
1141 AATCTGTGAAGATTAATCATTTCTCTCGAGACGTATATTGCACTGATCATATGACGA 1200
1201 GTGCTTCTATGGGAATTTACCTTGAACCTATTATGTTGTCCCTACTCTGTAGATTTAAT 1260
1201 GTGCTTCTATGGGAATTTACCTTGAACCTATTATGTTGTCCCTACTCTGTAGATTTAAT 1260
1261 TTTAGGAACCTCTCAGATACTTTTGAAGAGTACTGTAACTATATAGTCAACCCCTATAG 1320
1261 TTTAGGAACCTCTCAGATACTTTTGAAGAGTACTGTAACTATATAGTCAACCCCTATAG 1320
1321 TCACCTGGGCTTCAATTAAGAGATTCAGAACTGAATACCAACAGAAACAAACAGAACGA 1380
1321 TCACCTGGGCTTCAATTAAGAGATTCAGAACTGAATACCAACAGAAACAAACAGAACGA 1380
1381 CCAAAATTATGAATCATATAGTCAATAGGTTATCTCAATAGGGCTCATTTCACAATCTAG 1440
1381 CCAAAATTATGAATCATATAGTCAATAGGTTATCTCAATAGGGCTCATTTCACAATCTAG 1440
1441 GTGCATGTACCAAGTATATTTCTTGGACGCAACCGTAGTCAGATCGTACAAATACCATTAG 1500
1441 GTGCATGTACCAAGTATATTTCTTGGACGCAACCGTAGTCAGATCGTACAAATACCATTAG 1500
1501 TCAGATAGCATTAACAAATACCATTTGGTAAATCATTTCAACCTTAATTCAGTACTCTCT 1560
1501 TCAGATAGCATTAACAAATACCATTTGGTAAATCATTTCAACCTTAATTCAGTACTCTCT 1560
1561 GTAGTCAGTGGGCCAGGATTTACAGAGGGGATTAATCCGAACTAACCGTTAATGGTAGT 1620
1561 GTAGTCAGTGGGCCAGGATTTACAGAGGGGATTAATCCGAACTAACCGTTAATGGTAGT 1620
1621 GTACTAAGTATGGGTCTTAATTTAATAATAATCATCATTAACAGGGTATCGCGTAGAGTT 1680
1621 GTACTAAGTATGGGTCTTAATTTAATAATAATCATCATTAACAGGGTATCGCGTAGAGTT 1680
1681 CGTTATGCTGCTTCTCAACAAATGCTCGAGGTAACCTGTCGAGGGAGTACTACTTTT 1740
1681 CGTTATGCTGCTTCTCAACAAATGCTCGAGGTAACCTGTCGAGGGAGTACTACTTTT 1740
1741 GATCAAGGATTCCTTAGTACTATGATGCAAAATAGTCTTTGACATCTCAATCATTTAGA 1800
1741 GATCAAGGATTCCTTAGTACTATGATGCAAAATAGTCTTTGACATCTCAATCATTTAGA 1800
1801 TTTGCAAGATTTCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA 1860
1801 TTTGCAAGATTTCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA 1860
1861 AGTAATAATGCAAGGTAGACAAACGTTTCACTTTGATAAAATTTGAAATTCATTCCAATTACT 1920
1861 AGTAATAATGCAAGGTAGACAAACGTTTCACTTTGATAAAATTTGAAATTCATTCCAATTACT 1920
1921 GCAACCTTCGAAAGCAGAATACGATTTAGAAAGGCGCAAGAGCGGTGAATGCTCTGTTT 1980
1921 GCAACCTTCGAAAGCAGAATACGATTTAGAAAGGCGCAAGAGCGGTGAATGCTCTGTTT 1980
1981 ACTAATACGAATCCAAAGAGATTTGAAAACAGATGTGACAGATTTATCATATTGATCAAGTA 2040
1981 ACTAATACGAATCCAAAGAGATTTGAAAACAGATGTGACAGATTTATCATATTGATCAAGTA 2040
2041 TCCAAATTTAGTGGCGTGTATTATCGGATGAATTTCTGCTTAGATGAAAAGAGAGAAATTA 2100
2041 TCCAAATTTAGTGGCGTGTATTATCGGATGAATTTCTGCTTAGATGAAAAGAGAGAAATTA 2100
2101 GAGAAAGTGAATATATCGGAAACGACTCAGTGTAGTGAAGAAACCTTACTCCAAGATCCAAAC 2160
2101 GAGAAAGTGAATATATCGGAAACGACTCAGTGTAGTGAAGAAACCTTACTCCAAGATCCAAAC 2160
2161 TTCAATCATCATTAAGCAACGAGATCTCATATCTACTTAATGAGCAATCGAAATTTTACA 2220
2161 TTCAATCATCATTAAGCAACGAGATCTCATATCTACTTAATGAGCAATCGAAATTTTACA 2220

```
QY 2221 TCTATCCATGACAACTCTGAACATGGATGGTGGGGAAGTGAGAAACATTTACAAATCCAGGAA 2280
DB 2221 TCTATCCATGACAACTCTGAACATGGATGGTGGGGAAGTGAGAAACATTTACAAATCCAGGAA 2280
QY 2281 GGAAATGAGCTATTTAAAGAGAAATTTACGTACACTACCGGGGACCTTTTAAATGAGGTAT 2340
DB 2281 GGAAATGAGCTATTTAAAGAGAAATTTACGTACACTACCGGGGACCTTTTAAATGAGGTAT 2340
QY 2341 CCGAGCTATTTATATCAAAAATAGGAGTGGAAATTTAAAGCTTATACCTCGCTACCAA 2400
DB 2341 CCGAGCTATTTATATCAAAAATAGGAGTGGAAATTTAAAGCTTATACCTCGCTACCAA 2400
QY 2401 TTAAGAGGCTATTTAAAGAGTGTCAAGATTTAGAGATATATTTGATTCGTTTAAATCGG 2460
DB 2401 TTAAGAGGCTATTTAAAGAGTGTCAAGATTTAGAGATATATTTGATTCGTTTAAATCGG 2460
QY 2461 AAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATCGATCGATCGATCGATCGAT 2520
DB 2461 AAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATCGATCGATCGATCGATCGAT 2520
QY 2521 CCAATCGGAAGTGGGAGAACCGAATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 2580
DB 2521 CCAATCGGAAGTGGGAGAACCGAATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 2580
QY 2581 CTAGATTTGTTCTCTGAGAGATGGAGAAAATGTTGGCATCATTTCCCATCATTTCTCTTTG 2640
DB 2581 CTAGATTTGTTCTCTGAGAGATGGAGAAAATGTTGGCATCATTTCCCATCATTTCTCTTTG 2640
QY 2641 GATATTTGATTTGGATGACAGACTTTGATGAGAAATCTAGGCGTGGTGGTATTTCAAG 2700
DB 2641 GATATTTGATTTGGATGACAGACTTTGATGAGAAATCTAGGCGTGGTGGTATTTCAAG 2700
QY 2701 ATTAAGACGCGAGAGGTCATGCAAGACTAGGAAATCTGGAATTTATTGAGAGAAACCA 2760
DB 2701 ATTAAGACGCGAGAGGTCATGCAAGACTAGGAAATCTGGAATTTATTGAGAGAAACCA 2760
QY 2761 TTATTAGGAGAGCACTGCTCTGTTGAGAGAGCAGAGAAAATGAGAGAGACAAACGT 2820
DB 2761 TTATTAGGAGAGCACTGCTCTGTTGAGAGAGCAGAGAAAATGAGAGAGACAAACGT 2820
QY 2821 GAAAACTACAAATTTGAAACAAACAGATATATACAGAGGCAAAAGAGCTGTGGATGCT 2880
DB 2821 GAAAACTACAAATTTGAAACAAACAGATATATACAGAGGCAAAAGAGCTGTGGATGCT 2880
QY 2881 TTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
DB 2881 TTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCAT 2940
QY 2941 GCGGCAGATAAACTTGTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTGTTATC 3000
DB 2941 GCGGCAGATAAACTTGTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTGTTATC 3000
QY 3001 CCGGCTGTAATTCGGAATTTTGAAGAAATTAGAGGTCGCATTTATCACTGCAATCTCC 3060
DB 3001 CCGGCTGTAATTCGGAATTTTGAAGAAATTAGAGGTCGCATTTATCACTGCAATCTCC 3060
QY 3061 CTATACGATCGAGAAATGCTGTTAAATGTTGATTTTAAATGATTTGATTTAGCATGCTGG 3120
DB 3061 CTATACGATCGAGAAATGCTGTTAAATGTTGATTTTAAATGATTTGATTTAGCATGCTGG 3120
QY 3121 AATGTAAAGGCGATGTAGATGTCAACAGAGCCATCACCGTTCCTGCTTTGTTATCCCA 3180
DB 3121 AATGTAAAGGCGATGTAGATGTCAACAGAGCCATCACCGTTCCTGCTTTGTTATCCCA 3180
QY 3181 GAATGGGAGCAGAGTGTCAACAGCAGTTTCGCGTCTGTCGGGCGTGGCTATATCTTC 3240
DB 3181 GAATGGGAGCAGAGTGTCAACAGCAGTTTCGCGTCTGTCGGGCGTGGCTATATCTTC 3240
QY 3241 CGTGTACAGGCTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAATCGAG 3300
DB 3241 CGTGTACAGGCTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAATCGAG 3300
```

```
QY 3301 AACAAATACAGACGAACATAAAATTTAAAACTGTGAAGAGAGGAAGTGTATCCCAACGGAT 3360
DB 3301 AACAAATACAGACGAACATAAAATTTAAAACTGTGAAGAGAGGAAGTGTATCCCAACGGAT 3360
QY 3361 ACAGGAACGTGTAAATGATTTACTGCAACCAAGGTACAGCAGTATGTAATTTCCGTAAT 3420
DB 3361 ACAGGAACGTGTAAATGATTTACTGCAACCAAGGTACAGCAGTATGTAATTTCCGTAAT 3420
QY 3421 GCTCGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTTAAATTTACAAACCGACT 3480
DB 3421 GCTCGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTTAAATTTACAAACCGACT 3480
QY 3481 TATGAAGAGAAACGTATACAGATGTACGAAGAGATATCATTTGTGAATATGACAGAGGG 3540
DB 3481 TATGAAGAGAAACGTATACAGATGTACGAAGAGATATCATTTGTGAATATGACAGAGGG 3540
QY 3541 TATGTGAATTTACCAACCTACCACTGCTGTTATATGACAAAAGAAATTTAGAAATCTTCCA 3600
DB 3541 TATGTGAATTTACCAACCTACCACTGCTGTTATATGACAAAAGAAATTTAGAAATCTTCCA 3600
QY 3601 GAAACCGATATAGGTATGATTTGAGATTGGAGAAACGGAAGGAAAGTTTATTTGTAGACAGC 3660
DB 3601 GAAACCGATATAGGTATGATTTGAGATTGGAGAAACGGAAGGAAAGTTTATTTGTAGACAGC 3660
QY 3661 GTGGAATTTACTCTCTTATGGAGGAATAG 3687
DB 3661 GTGGAATTTACTCTCTTATGGAGGAATAG 3687
```

RESULT 2

```
US-10-428-961-37
; Sequence 37, Application US/10428961
; Publication No. US2003023711A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Ruper, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE REFERENCE: Polynucleotides, Compositions, and Methods of Use (Amended)
; TITLE OF INVENTION: MECO201--1
; CURRENT APPLICATION NUMBER: US/10/428,961
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 4173
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3687)
US-10-428-961-37
```

```
Query Match 99.4%; Score 3666.2; DB 15; Length 4173;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3674; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 TTGACTTCAATAGAAAAATGAGAAATTAATTAATGCTTTTATCGATCCAGCTGTA 60
DB 1 TTGACTTCAATAGAAAAATGAGAAATTAATTAATGCTTTTATCGATCCAGCTGTA 60
QY 61 TCGAATCATTTCCACACAAATGATCTATCACAGATGCTCGTATTTGAGGATTTCTTTGTGT 120
DB 61 TCGAATCATTTCCACACAAATGATCTATCACAGATGCTCGTATTTGAGGATTTCTTTGTGT 120
QY 121 ATAGCCGAGGGGAAATTAATATCAATTTGTTAGCGCATCAACAGTCCAAACCGGTATT 180
DB 121 ATAGCCGAGGGGAAATTAATATCAATTTGTTAGCGCATCAACAGTCCAAACCGGTATT 180
```

Db 121 ATAGCCGAGGGAATAATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATT 180
Qy 181 AACATAGCTGTAGAAATCTAGGTGTATTAGGGGTACCGTTTGTGACAAATAGCTAGT 240
Db 181 AACATAGCTGTAGAAATCTAGGTGTATTAGGGGTACCGTTTGTGACAAATAGCTAGT 240
Qy 241 TTTTATAGTTTCTTGTGGTGAATATTAGGCCCGCGCAGAGATCAGTGGGAATTTTC 300
Db 241 TTTTATAGTTTCTTGTGGTGAATATTAGGCCCGCGCAGAGATCAGTGGGAATTTTC 300
Qy 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAAATGCTAGGAATACGGCA 360
Db 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAAATGCTAGGAATACGGCA 360
Qy 361 CTTCCTCGATTACAGGTTTAGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
Db 361 CTTCCTCGATTACAGGTTTAGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
Qy 421 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCCAATATATAGCC 480
Db 421 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCCAATATATAGCC 480
Qy 481 TTAGAACTTGATTTCTTAATGCGATCGCGCTTTTCGCAANTTAGAAAACCAAGAGTTCCA 540
Db 481 TTAGAACTTGATTTCTTAATGCGATCGCGCTTTTCGCAANTTAGAAAACCAAGAGTTCCA 540
Qy 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATATTATTGAGAGATGCGCTCT 600
Db 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATATTATTGAGAGATGCGCTCT 600
Qy 601 CTTTTTGGTAGTAATTTGGGCTTACATCGCAGGAATTCACCGTTATTATTAGCGGCCAA 660
Db 601 CTTTTTGGTAGTAATTTGGGCTTACATCGCAGGAATTCACCGTTATTATTAGCGGCCAA 660
Qy 661 GTGGAAACAAACGAGAGATTATTCGCACTATGCGTAGAATGGTATGATTAATACAGGCTCTAAAT 720
Db 661 GTGGAAACAAACGAGAGATTATTCGCACTATGCGTAGAATGGTATGATTAATACAGGCTCTAAAT 720
Qy 721 AGCTTGAGAGGGAACAAATGCGCAAGTTGGGTGCTTATTAATCAATTCGCTAGAGATCTA 780
Db 721 AGCTTGAGAGGGAACAAATGCGCAAGTTGGGTGCTTATTAATCAATTCGCTAGAGATCTA 780
Qy 781 ACGTTAGGGGTAATAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 840
Db 781 ACGTTAGGGGTAATAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCACTTATCCA 840
Qy 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCCCAATTTGAGCAACAGGG 900
Db 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCCCAATTTGAGCAACAGGG 900
Qy 901 GTAAATATGCAAGTATGAAATGCTATATAATAATGCACTTCGTTTCCGCTATAGAG 960
Db 901 GTAAATATGCAAGTATGAAATGCTATATAATAATGCACTTCGTTTCCGCTATAGAG 960
Qy 961 ACTGCGGTTATCCGAAGCCGCACTACTTGATTTTCTAGAACAACTTACAAATTTTTCAG 1020
Db 961 ACTGCGGTTATCCGAAGCCGCACTACTTGATTTTCTAGAACAACTTACAAATTTTTCAG 1020
Qy 1021 ACTTCATCAGATGGAGTGCTACTAGGCATATGACCTTACTGCGGGGGGACACAATTCAA 1080
Db 1021 ACTTCATCAGATGGAGTGCTACTAGGCATATGACCTTACTGCGGGGGGACACAATTCAA 1080
Qy 1081 TCTCGGCCAATAGAGCGGATTAATACTCAACGATGGGTCTACCAATCTCTATT 1140
Db 1081 TCTCGGCCAATAGAGCGGATTAATACTCAACGATGGGTCTACCAATCTCTATT 1140
Qy 1141 AATCCTGTAAAGATTATCTTCTCTCGAGAGTATATTGAGTGAATCATATGACGGA 1200
Db 1141 AATCCTGTAAAGATTATCTTCTCTCGAGAGTATATTGAGTGAATCATATGACGGA 1200
Qy 1201 GTGCTTCTATGGGGAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTTAAT 1260
Db 1201 GTGCTTCTATGGGGAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTTAAT 1260

Qy 1261 TTTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG 1320
Db 1261 TTTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG 1320
Qy 1321 TCACCTGGGCTTCAATTAAGAAATTCAGAAAATGAAATACCAACGAGAAAACAGAACGA 1380
Db 1321 TCACCTGGGCTTCAATTAAGAAATTCAGAAAATGAAATACCAACGAGAAAACAGAACGA 1380
Qy 1381 CCAAAATTAAGAAATCATATAGTCTATAGTCTCACAATAGGGCTCATTTACAAATCTAGG 1440
Db 1381 CCAAAATTAAGAAATCATATAGTCTATAGTCTCACAATAGGGCTCATTTACAAATCTAGG 1440
Qy 1441 GTGCATGTACCAAGTATATTCTTTGGACGACCGTAGTGCAGATCGTACAAATACCAATAGT 1500
Db 1441 GTGCATGTACCAAGTATATTCTTTGGACGACCGTAGTGCAGATCGTACAAATACCAATAGT 1500
Qy 1501 TCAGATAGCATTAACACAAATACCATTTGGTAAATCATTTCAACCTTTAATTCAGGTACTCT 1560
Db 1501 TCAGATAGCATTAACACAAATACCATTTGGTAAATCATTTCAACCTTTAATTCAGGTACTCT 1560
Qy 1561 GTAGTCAGTGGGCCAGGATTTTACAGGAGGGGATATATCCGAACTAAACGTTAATGCTAGT 1620
Db 1561 GTAGTCAGTGGGCCAGGATTTTACAGGAGGGGATATATCCGAACTAAACGTTAATGCTAGT 1620
Qy 1621 GTACTAAGTATGGGCTTTAATTTAATAATCATCATTTACAGCGGTATCGCGTAGAGTT 1680
Db 1621 GTACTAAGTATGGGCTTTAATTTAATAATCATCATTTACAGCGGTATCGCGTAGAGTT 1680
Qy 1681 CGTTATGCTGCTTCACAACTGCTGAGGGTAACTGTCGGAGGGAGTACTACTTTT 1740
Db 1681 CGTTATGCTGCTTCACAACTGCTGAGGGTAACTGTCGGAGGGAGTACTACTTTT 1740
Qy 1741 GATCAAGGATTCCTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCATTTAGA 1800
Db 1741 GATCAAGGATTCCTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAATCATTTAGA 1800
Qy 1801 TTTGCAAGATTTTCTGTAGGTATTAGTGATCTGCGAGTCAAACTGCTGGAATAAGTATA 1860
Db 1801 TTTGCAAGATTTTCTGTAGGTATTAGTGATCTGCGAGTCAAACTGCTGGAATAAGTATA 1860
Qy 1861 AGTAATATGCAAGTATGCAAACTGTTTCACTTTGATAAATTTGAATTCATTTCCAATTACT 1920
Db 1861 AGTAATATGCAAGTATGCAAACTGTTTCACTTTGATAAATTTGAATTCATTTCCAATTACT 1920
Qy 1921 GCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGAGCGGTGAATGCTCTGTTT 1980
Db 1921 GCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGAGCGGTGAATGCTCTGTTT 1980
Qy 1981 ACTAATACGAATCCAAAGAGATTGAAAACAGATGTGACAGATTATCATATTGATCAAGTA 2040
Db 1981 ACTAATACGAATCCAAAGAGATTGAAAACAGATGTGACAGATTATCATATTGATCAAGTA 2040
Qy 2041 TCCAAATTTAGTGGCGTGTATTCCGATGAATTTCTGTTAGATGAAAAGAGAGAAATTACTT 2100
Db 2041 TCCAAATTTAGTGGCGTGTATTCCGATGAATTTCTGTTAGATGAAAAGAGAGAAATTACTT 2100
Qy 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGAATGAAGAAAACCTTACTCCAAGATCCAAC 2160
Db 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGAATGAAGAAAACCTTACTCCAAGATCCAAC 2160
Qy 2161 TTCATCATCAATCAATAGCAACCGAGCTTCAATCTACTATGATGAGCAATCGAATTTTACA 2220
Db 2161 TTCATCATCAATCAATAGCAACCGAGCTTCAATCTACTATGATGAGCAATCGAATTTTACA 2220
Qy 2221 TCTATCCAATGAACATCTGAAACATGGATGGTGGGGAAGTGAGAACATTAACAATCCAGAA 2280
Db 2221 TCTATCCAATGAACATCTGAAACATGGATGGTGGGGAAGTGAGAACATTAACAATCCAGAA 2280
Qy 2281 GGAATGAGCTATTAAAGAGAAATTAAGTCACTACCGGGGACTTTTAAATGAGTGTAT 2340
Db 2281 GGAATGAGCTATTAAAGAGAAATTAAGTCACTACCGGGGACTTTTAAATGAGTGTAT 2340

```
Qy 2341 CCGACGATTTATATCAAAAAATAGGAGTCGGAATTTAAAGCTTATACCTCGCTACCA 2400
Db 2341 CCGACGATTTATATCAAAAAATAGGAGTCGGAATTTAAAGCTTATACCTCGCTACCA 2400
Qy 2401 TTAAAGGGTATATTTGAAGATAGTCAGATTTAGAGATATATTTGATTCGTTTAAATGCG 2460
Db 2401 TTAAAGGGTATATTTGAAGATAGTCAGATTTAGAGATATATTTGATTCGTTTAAATGCG 2460
Qy 2461 AAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGTCGGCGCTTTCAGTTGAAAGC 2520
Db 2461 AAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGTCGGCGCTTTCAGTTGAAAGC 2520
Qy 2521 CCAATCGAAGGTCGGAGAACCGAATCGATCGCACACATTTTGAATGAAATCCCTGAT 2580
Db 2521 CCAATCGAAGGTCGGAGAACCGAATCGATCGCACACATTTTGAATGAAATCCCTGAT 2580
Qy 2581 CTAGATTTCTCTCGAGAGATGGGAGAAATGTCGCATCATTTCCCATCATTTCTCTTTG 2640
Db 2581 CTAGATTTCTCTCGAGAGATGGGAGAAATGTCGCATCATTTCCCATCATTTCTCTTTG 2640
Qy 2641 GATATTGATATTGATGCACAGATTTGATGAGATCTAGGCGTGTGGTGGTATTCAAG 2700
Db 2641 GATATTGATATTGATGCACAGATTTGATGAGATCTAGGCGTGTGGTGGTATTCAAG 2700
Qy 2701 ATTAAGACCGAGGATGTCATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAAACCA 2760
Db 2701 ATTAAGACCGAGGATGTCATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAAACCA 2760
Qy 2761 TTATTAGGAGAGCACTGCTCTCGTGTGAGAGAGCAGAGAAATGTCGAGACAAAGCT 2820
Db 2761 TTATTAGGAGAGCACTGCTCTCGTGTGAGAGAGCAGAGAAATGTCGAGACAAAGCT 2820
Qy 2821 GAAAACTACAAATTTGGAACCAAAACGAGTATATACAGAGCAGGCAAGAGCTGTGGATGCT 2880
Db 2821 GAAAACTACAAATTTGGAACCAAAACGAGTATATACAGAGCAGGCAAGAGCTGTGGATGCT 2880
Qy 2881 TTATTGATGATTTCTCAATATAATAGATTACAGCGGATACAAACATTTGGCATGATTCAT 2940
Db 2881 TTATTGATGATTTCTCAATATAATAGATTACAGCGGATACAAACATTTGGCATGATTCAT 2940
Qy 2941 GCGGAGATTAACCTGTTTCATGCAATTCGAGAGCTTATCTGTGAGATTTATCTGTTATC 3000
Db 2941 GCGGAGATTAACCTGTTTCATGCAATTCGAGAGCTTATCTGTGAGATTTATCTGTTATC 3000
Qy 3001 CCGGCTGTAATCGGAAATTTTGAAGATTTAGAGCTGCAATTTACCTGCAATCTCC 3060
Db 3001 CCGGCTGTAATCGGAAATTTTGAAGATTTAGAGCTGCAATTTACCTGCAATCTCC 3060
Qy 3061 CTATACGATCGAGAAATGTCGTTAAAAATGGTGATTTTAAATATGATGATGATGCTGG 3120
Db 3061 CTATACGATCGAGAAATGTCGTTAAAAATGGTGATTTTAAATATGATGATGATGCTGG 3120
Qy 3121 AATGTAAGGGCATGTAGATGTACACAGAGCCATCACCGTCTGTCTGCTTGTATCCCA 3180
Db 3121 AATGTAAGGGCATGTAGATGTACACAGAGCCATCACCGTCTGTCTGCTTGTATCCCA 3180
Qy 3181 GAATGGGAAGCAGAAAGTGTCAAGAGCTGTCGCGTCTGTCCGGCGGTGGCTATATCTC 3240
Db 3181 GAATGGGAAGCAGAAAGTGTCAAGAGCTGTCGCGTCTGTCCGGCGGTGGCTATATCTC 3240
Qy 3241 CGTGTACAGCGTACAAAGAGGGATGTGGAGAGGGTTGTGTAAACGATCCATGAAATCGAG 3300
Db 3241 CGTGTACAGCGTACAAAGAGGGATGTGGAGAGGGTTGTGTAAACGATCCATGAAATCGAG 3300
Qy 3301 AACATACAGACGAACTTAAATTTTAAACCTGTGAAGAGAGGAGTGTATCCAAACGAT 3360
Db 3301 AACATACAGACGAACTTAAATTTTAAACCTGTGAAGAGAGGAGTGTATCCAAACGAT 3360
Qy 3361 ACAGGAAGCTGTATGATTTACTGCAACCAAGCTACAGAGTGTGATTTCCCGTAAT 3420
Db 3361 ACAGGAAGCTGTATGATTTACTGCAACCAAGCTACAGAGTGTGATTTCCCGTAAT 3420
Qy 3421 GCTGGATATGAGGATGATATGAAGTTGATACAGCATCTGTTAAATTAACAAACCGACT 3480
```

```
Db 3421 GCTGGATATGAGGATGATATGAAGTTGATACAGCATCTGTTAAATTAACAAACCGACT 3480
Qy 3481 TATCAAGAGAAAGCTATACAGATGTACGAGATGAATCAATCTGATATGATACAGAGG 3540
Db 3481 TATCAAGAGAAAGCTATACAGATGTACGAGATGAATCAATCTGATATGATACAGAGG 3540
Qy 3541 TATGTGAATTTATCCACCACTACCACTGCTGTTATATGACAAAGAAATTAGAATATCTCCCA 3600
Db 3541 TATGTGAATTTATCCACCACTACCACTGCTGTTATATGACAAAGAAATTAGAATATCTCCCA 3600
Qy 3601 GAAACCGATTAAGTATGATGATTTGAGATTTGGAAGAAACGGAAGGAATTTATTTGATAGACAGC 3660
Db 3601 GAAACCGATTAAGTATGATGATTTGAGATTTGGAAGAAACGGAAGGAATTTATTTGATAGACAGC 3660
Qy 3661 GTGGAATTTACTCTCTTATGAGGAGATAG 3687
Db 3661 ATAGAATTTACTCTCTTATGGAAGATAG 3687

RESULT 3
US-10-428-961-62
; Sequence 62, Application US/10428961
; Publication No. US20030237111A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Ruper, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE REFERENCE: MCO201-1
; CURRENT APPLICATION NUMBER: US/10/428,961
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-428-961-62

Query Match 88.4%; Score 3260.2; DB 15; Length 3684;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

Qy 1 TTGACTTCAATAGGAAAAATGGAATGAATTAATTAATGCTTTATCGATTCAGCTGTA 60
Db 1 TTGACTTCAATAGGAAAAATGGAATGAATTAATTAATGCTTTATCGATTCAGCTGTA 60
Qy 61 TCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTTGAGGATCTTTTGGT 120
Db 61 TCGAATCATTTCCGACAAATGAATCTATCAACCGATGCTCGTATTTGAGGATGCTTTGGT 120
Qy 121 ATAGCCGAGGGGAATAATATCAATCCACTTTGTAGCCATCAACAGTCCAAACCGGTATT 180
Db 121 ATAGCCGAGGGGAATAATATCAATCCACTTTGTAGCCATCAACAGTCCAAACCGGTATT 180
Qy 181 AACATAGCTGGTAGAATCTAGGTGATTTAGGCGTACGTTTGTGCAAAATAGCTAGT 240
Db 181 AACATAGCTGGTAGAATCTAGGTGATTTAGGCGTACGTTTGTGCAAAATAGCTAGT 240
Qy 241 TTTTATAGTTTCTTTGTTGGTGAATTTATGCGCGGCGGAGAGATCAGTGGGAAATTTTC 300
Db 241 TTTTATAGTTTCTTTGTTGGTGAATTTATGCGCGGCGGAGAGATCCTTGGGAAATTTTC 300
Qy 301 CTAGAACATGTCGAAACCACTTATAAATCAACAAATAACAGAAATGCTTAGGAATACGCA 360
```

Db 301 CTAGAACATGTCGAACATCTTTAAGACAAACAAGTACAGAAAATACTAGGATACGGCT 360
Qy 361 CTTGCTCGATTACAGGTTTAGAGATTCTCTTTAGAGCCCTATCAAAGTCACCTTGAAGAT 420
Db 361 CTTGCTCGATTACAGGTTTAGGAAATTCCTTTAGAGCCCTATCAACAGTCACCTTGAAGAT 420
Qy 421 TGGCTAGAAAACCGTGATGATCAAGAACGAGAAAGTGTTCTTTATACCCAAATATATAGCC 480
Db 421 TGGCTAGAAAACCGTGATGATCAAGAACGAGAAAGTGTTCTTTATACCCAAATATATAGCC 480
Qy 481 TTAGAACTTGATTTTCTTAATCGATGCGCGCTTTTCGCAATTAGAAAACCAAGAAGTTCCA 540
Db 481 TTAGAACTTGATTTTCTTAATCGATGCGCGCTTTTCGCAATTAGAAAACCAAGAAGTTCCA 540
Qy 541 TTATTAATGATATATGCTCAAGCTGCBAATTTACACTTATATATTTAGAGAGATGCTCT 600
Db 541 TTATTAATGATATATGCTCAAGCTGCBAATTTTACACTTATATATTTAGAGAGATGCTCT 600
Qy 601 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATCAACGTTTATTATGAGCGCCAA 660
Db 601 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATCAACGTTTATTATGAGCGCCAA 660
Qy 661 GTGGAAACAAAGAGAGATTATCCGACTATTCGCTAGAAATGGGTATTAATACAGTCTAAAT 720
Db 661 GTGGAAACAAAGAGAGATTATCTGATTTATTTGGCAAGATGGTATTAATACGGGTTTAAAT 720
Qy 721 AGCTTCGAGAGGCAAAATCGCAAGTTCGGTTCGTTTAAATCAATTCGCTAGAGATCTA 780
Db 721 AATTTGAGAGGCAAAATCGTGAAGTTCGGTTCGATATAATCAATTCGCTAGAGATCTA 780
Qy 781 ACCTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 840
Db 781 ACCTTAGGAGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACGCGTGTATCCA 840
Qy 841 ATAAATACGAGTGCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTCGAGCAACAGGG 900
Db 841 ATGAATACCAAGTGCTCAATTTAAACAGAGAAATTTATACAGATCCAAATTTGGGAGAACAAAT 900
Qy 901 GTAAAT- ----ATGCAAGTATGAATTCGTATAATAATAATGAACCTTCGTTTCCGCT 954
Db 901 GCACCTTCAGGATTTGCAAGTACGAATTCGTTTAAATAATGAACCACTTCGTTTCCGCC 960
Qy 955 ATAGAGATCGCGTTATCCGAAGCCCGCATCTACTTTGATTTTCTAGAACAACTTACAAT 1014
Db 961 ATAGAGGCTGCGGTTATTAGGCTCCGCATCTACTTTGATTTTCCGAACAGCTTACAAT 1020
Qy 1015 TTATGACATTCATCAGATGGAGTCTACTAGGCATATGACTTACTGGCGGGGCAACA 1074
Db 1021 TTAGCGGTATTAGTCCGATGGAGTAACTCAATATATGAATTTACTGGGTGGGACATAGA 1080
Qy 1075 ATTCAATCTCGGCCAATAGAGGCGGATTAATAATACCTCAACGATGGGTCTACCAATCT 1134
Db 1081 CTTGNAATCGGACATATAGGGGGTCATTTAGTACCTCGACACACGGAATACCAATCT 1140
Qy 1135 TCTATTAACTCTGTAGATTATCAATTCCTCTCGAGACGTATATTTGGACTGAATCATAT 1194
Db 1141 TCTATTAACTCTGTAACTTACAGTTTCAATCTCGAGACGTTTATAGAACAGAAATCATTT 1200
Qy 1195 GCAGGAGTCTTCTATGGGGAATTTACCTTGAACCTTATCATGGTTCCTACTGTAGA 1254
Db 1201 GCAGGATTAATAT- ----ACTTTAACTACTCTCTGTAATGGAGTACCTTGGGCTAGA 1254
Qy 1255 TTTAATTTTAGAACCTCTCAGAACTCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCC 1314
Db 1255 TTTAATTTGGAGAAATCCCTCGAATTCCTTT- --AGAGGTAGCTTCTCTATATAGGG 1311
Qy 1315 TATGACTCACCTGGGCTTCAATTTAAAGATTCGAAACCTGAATACCAACGAGAAACA 1374
Db 1312 TATACTGGAGTGGGGAACAACCTATTTGATTCAGAACTGAATTTACCAACGAGAAACA 1371
Qy 1375 GAACGACCAAAATTTATGAATCATATAGTCTATAGTTATCTCACATAGGGCTCATTTCAAA 1434
Db 1372 GAACGACCAAAATTTATGAATCTTACAGTCTATAGATTTCTAAATATAGACTAATATACGGA 1431

Qy 1435 TCTAGGGTCGATGTACCAGTATATTCTTGGAGCGACCCGCTAGTGCAGATCGTACAAATACC 1494
Db 1432 AACACCTTTGGAGAGCACCAGTATATTCTTGGAGCGACCCGCTAGTGCAGATCGTACAAATACC 1491
Qy 1495 ATTTAGTTCCAGATAGATACACAAATACCATTGGTAAATCAATCAACCTTAAATTCAGGT 1554
Db 1492 ATTTAGTTCCAGATAGATACACAAATACCAATTTGGTAAATCAATCAACCTTAAATTCAGGT 1551
Qy 1555 ACCTCTGTAGTCAGTGGCCAGGATTTTACAGAGGGGATATTAATCCGAACCTAACGTTAAT 1614
Db 1552 ACCTCTGTAGTCAGTGGCCAGGATTTTACAGAGGGGATATTAATCCGAACCTAACGTTAAT 1611
Qy 1615 GGTAGTGTACTAAGTATGGGCTCTTAATTTTAAATATACATCAATACAGCGGTATCCCGTG 1674
Db 1612 GGTAGTGTACTAAGTATGGGCTCTTAATTTTAAATATACATCAATACAGCGGTATCCCGTG 1671
Qy 1675 AGAGTTTCGTTATGCTGCTTCTCAAAATGGTTCCTGAGGGTAACTGTGCGAGGGAGTACT 1734
Db 1672 AGAGTTTCGTTATGCTGCTTCTCAAAATGGTTCCTGAGGGTAACTGTGCGAGGGAGTACT 1731
Qy 1735 ACTTTTGATCAAGGATTCCTTAGTACTATAGTGCAAATGAGTCTTTTGACATCTCAATCA 1794
Db 1732 ACTTTTGATCAAGGATTCCTTAGTACTATAGTGCAAATGAGTCTTTTGACATCTCAATCA 1791
Qy 1795 TTTAGATTTGCAGAAATTTCCCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAAAT 1854
Db 1792 TTTAGATTTGCAGAAATTTCCCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAAAT 1851
Qy 1855 AGTATAAGTAAATTAATGCAGGTAGACAAAAGTTTCACTTTTGATAAAAATTTGAAATTCATTC 1914
Db 1852 AGTATAAGTAAATTAATGCAGGTAGACAAAAGTTTCACTTTTGATAAAAATTTGAAATTCATTC 1911
Qy 1915 ATTACTGCMACTTCGAAAGCAGAAATACGATTTAGAAAAGGCGCCAAAGCGCGTGAATGCT 1974
Db 1912 ATTACTGCMACTTCGAAAGCAGAAATATGATTTAGAAAAGGCGCCAAAGCGCGTGAATGCT 1971
Qy 1975 CTGTTTACTTAATPACCAATCCCAAGAGATTGAAAACAGATGTGACAGATTTATCATATTGAT 2034
Db 1972 CTGTTTACTTAATPACCAATCCCAAGAGGTTGAAAACAGGTGTGACAGATTTATCATATTGAT 2031
Qy 2035 CAAATATCCAAATTTAGTGGCGTGTATTCGAGTGAATTTCTGCTTAGATGAAAAGAGAGAA 2094
Db 2032 GAAGTATCCAAATTTAGTGGCGTGTATTCGAGTGAATTTCTGCTTGAATGAAAAGAGAGAA 2091
Qy 2095 TTACTTTGAAAAGTGAATATGCGAAAACGACTCAGTGAATGAAAAGAACTTACTTCCAAAT 2154
Db 2092 TTACTTTGAAAAGTGAATATGCGAAAACGACTCAGTGAATGAAAAGAACTTACTTCCAAAT 2151
Qy 2155 CCAAACTTCAATCCATCAATTAAGCAACGAGACTTTCATATCTACTAATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCAATCCATCAATTAAGCAACGAGACTTTCATATCTAATTAATGAGCAATCGAAT 2211
Qy 2215 TTCACATCTATCCATGAAACAACTCTGAAACATGGAATGGTGGGGAAGTGAGAACATTTACAATC 2274
Db 2212 TTCACATCTATCCATGAAACAACTCTGAAACATGGAATGGTGGGGAAGTGAGAACATTTACAATC 2271
Qy 2275 CAGGAAGGAAATGAGTATTTAAAGAGAAATTAACGTCACTACACCGGGGACTTTTAAATGAG 2334
Db 2272 CAGGAAGGAAATGAGTATTTAAAGAGAAATTAACGTCACTACACCGGGGACTTTTAAATGAG 2331
Qy 2335 TGTTATCCGAGCTATTTTATATCAAAAAATAGGAGAGTCCGGAATTTAAAAGCTTTTAACTCGC 2394
Db 2332 TGTTATCCGAGCTATTTTATATCAAAAAATAGGAGGCGGAATTTAAAAGCTTTTAACTCGC 2391
Qy 2395 TACCAATTAAGAGGGTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2454
Db 2392 TACCAATTAAGTGGCTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2451
Qy 2455 AATGCGAAACATGAACCAATTTGGATTCAGGTACCGAGTCCGATTCGCGCTTTTCAGTT 2514
Db 2452 AATGCGAAACATGAACCAATTTGGATTCAGGTACCGAGTCCGATTCGCGCTTTTCAGTT 2511

QY 2515 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGCACCAATTTTGAATGGAAT 2574
DB 2512 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGCACCAATTTTGAATGGAAT 2571
QY 2575 CTTGATCTAGATTTCTCTGCAGAGATGGAGAAAATGTGGCATCATTTCCATCATTTT 2634
DB 2572 CTTGATCTAGATTTCTCTGCAGAGATGGAGAAAATGTGGCATCATTTTCCATCATTTT 2631
QY 2635 TCTTTGGATATTGATATTGGATGCAAGATTTGCATGAGAAATCTAGGCGTGTGGTGGA 2694
DB 2632 TCTTTGGATATTGATTTGGATGCATAGCTTGCATGAGAACCTAGGCGTGTGGTGGA 2691
QY 2695 TTCAAGATTAGACCGCAGGAAGGTGCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAG 2754
DB 2692 TTCAAGATTAGACCGCAGGAAGGTGCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAG 2751
QY 2755 AAACATTTATTAGGAGAGCACTGTCTGTGTGAAGAGAGCAGAGAAAATGGAGAC 2814
DB 2752 AAACATTTATTAGGAGAGCACTGTCTGTGTGAAGAGAGCAGAGAAAATGGAGAC 2811
QY 2815 AAACGTGAAAACTACAATTGGAAAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2874
DB 2812 AAACGTGAAAACTACAATTGGAAAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2871
QY 2875 GATGCTTTATTGTAGATTCTCAATATATAGATTACAAGCGGATACAAACATTTGGCATG 2934
DB 2872 GATGCTTTATTGTAGATTCTCAATATATAGATTACAAGCGGATACAAACATTTGGCATG 2931
QY 2935 ATTGATCGGCAGATAAATCTGTTCATCGAATTCGAGAGGCTTATCTGTCAAGATTATCT 2994
DB 2932 ATTGATCGGCAGATAAATCTGTTCATCGAATTCGAGAGGCTTATCTGTCAAGATTATCT 2991
QY 2995 GTTATCCCGGTGTAATTCGGAATTTTGAAGAAATTTAGAAAGTTCGCAATTTCACTGCA 3054
DB 2992 GTTATCCCGGTGTAATTCGGAATTTTGAAGAAATTTAGAAAGTTCGCAATTTCACTGCA 3051
QY 3055 ATCTCCCTATACGATCGAGAAATGTCGTTAAAAATGTTGATTTTAAATGGAATGATGCA 3114
DB 3052 ATCTCCCTATACGATCGAGAAATGTCGTTAAAAATGTTGATTTTAAATGGAATGATGCA 3111
QY 3115 TGCTGGATGTAAAGGCGATGTAGATGTACACAGAGCCATCAGCTTCTGCTGTTGTT 3174
DB 3112 TGCTGGATGTAAAGGCGATGTAGATGTACACAGAGCCATCAGCTTCTGCTGTTGTT 3171
QY 3175 ATCCAGAAATGGGAAGCAGAAAGTGTCAAGCAGTTTCGCGTCTGTCCGGGCGTGGCTAT 3234
DB 3172 ATCCAGAAATGGGAAGCAGAAAGTGTCAAGCAGTTTCGCGTCTGTCCGGGCGTGGCTAT 3231
QY 3235 ATCTCCGTGTACAGCGGTACAAAGAGGGATATGGAGAGGGTGTGTAAACGATCCATGAA 3294
DB 3232 ATCTCCGTGTACAGCGGTACAAAGAGGGATATGGAGAGGGTGTGTAAACGATCCATGAA 3291
QY 3295 ATCGAGAAATACAGAGCACTAAATTTTAAATCTGTGAAGAGAGGAAGTGTATCCA 3354
DB 3292 ATCGAGAAATACAGAGCACTAAATTTTAAATCTGTGAAGAGAGGAAGTGTATCCA 3351
QY 3355 ACGGATACAGAACTGTGAATGATTATCTACACCAAGGTACAGCAGTATGTAATTTCC 3414
DB 3352 ACGGATACAGAACTGTGAATGATTATCTACACCAAGGTACAGCAGTATGTAATTTCC 3411
QY 3415 CGTAATGCTGGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTTAAATTTACAAA 3474
DB 3412 CGTAATGCTGGATATGAGGATGCATATGAAGTTGATCTACAGCATCTGTTAAATTTACAAA 3471
QY 3475 CCGACTTATGAAGAGAAACGTATACAGATGTACGAAGAGATAATCATTTGTAATATGAC 3534
DB 3472 CCGACTTATGAAGAGAAACGTATACAGATGTACGAAGAGATAATCATTTGTAATATGAC 3531
QY 3535 AGAGGGTATGGAATTTATCCACCACTACAGCTGTTATATGACAAAAGAAATTTAGATAC 3594
DB 3532 AGAGGGTATGGAATTTATCCACCACTACAGCTGTTATATGACAAAAGAAATTTAGATAC 3591
QY 3595 TTCCAGAAAACCGATAAGGTATGGATTGGAGAAACCGAAGGGAAGTTTATTGTA 3654

DB 3592 TTCCAGAAAACCGATAAGGTATGGATTGGAGAAACCGAAGGGAAGTTTATTGTA 3651
QY 3655 GACAGCGTGGAAATTAATCTCTTATGAGGAATAG 3687
DB 3652 GACAGCGTGGAAATTAATCTCTTATGAGGAATAG 3684

RESULT 4
US-10-809-953-9
; Sequence 9, Application US/10809953
; Publication No. US20040181825A1
; GENERAL INFORMATION:
; APPLICANT: Van Mellaert, Herman
; APPLICANT: Botterman, Johan
; APPLICANT: Van Rie, Jeroen
; APPLICANT: Joes, Henk
; TITLE OF INVENTION: RECOMBINANT PLANT EXPRESSING NON-COMPETITIVELY BINDING Bt INSECTI
; TITLE OF INVENTION: CRYSTAL PROTEINS
; FILE REFERENCE: 021565-078
; CURRENT APPLICATION NUMBER: US/10/809,953
; PRIORITY FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US/09/661,016
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: PCT/EP90/00905
; PRIOR FILING DATE: 1990-05-30
; PRIOR APPLICATION NUMBER: GB 89401499.2
; PRIOR FILING DATE: 1989-05-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3687
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(3687)
US-10-809-953-9

Query Match 85.0%; Score 3133.6; DB 17; Length 3687;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 299; Indels 30; Gaps 3;

QY 1 TTGACTTCAAAATAGAAAAATGAGAATGAAAATATAAATGCTTTATCGATTCACAGCTGTA 60
DB 1 TTGACTTCAAAATAGAAAAATGAGAATGAAAATATAAATGCTTTATCGATTCACAGCTGTA 45
QY 61 TCGAATCATTCACACAAAATGGATCTATCACAGATGCTCGTATTTGAGGATTTCTTTGTGT 120
DB 46 TCGAATCATTCGCGCACAAATGGATCTATTAACAGATGCTCGTATTTGAGGATAGCTTTGT 105
QY 121 ATAGCCGAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACCGGTATT 180
DB 106 ATAGCCGAGGGGAACAAATATTTGATCCATTTGTTAGCGCATCAACAGTCCAAACCGGTATT 165
QY 181 AACATAGCTGTAGAACTAGGTGTATTAGCGGTACCGTTTGTGGCAAAATAGCTAGT 240
DB 166 AACATAGCTGTAGAACTAGGTGTATTAGCGGTATTTGGCGTATCGTTTGTGGCAAAATAGCTAGT 225
QY 241 TTTTATAGTTTCTTTGTTGGTAAATTTATGGCCCCCGCGCAGAGATCACTAGTGGGAAATTTTC 300
DB 226 TTTTATAGTTTCTTTGTTGGTAAATTTATGGCCCCCGCGCAGAGATCACTAGTGGGAAATTTTC 285
QY 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATACAGAAAATGCTAGGAATACGGA 360
DB 286 CTAGAACATGTCGAACAACTTATAAATCAACAAATACAGAAAATGCTAGGAATACGGA 345
QY 361 CTTCTCTGATTTACAGGTTTAGGAGATTCTTTTAGAGCTTCAACAGTCACTTGAAGAT 420
DB 346 CTTCTCTGATTTACAGGTTTAGGAGATTCTTTTAGAGCTTCAACAGTCACTTGAAGAT 405
QY 421 TGGCTAGAAAACCGGTATGATGCAAGAACCGAAGAGTGTCTTTATACCCAAATATATAGCC 480

Db 406 TGGCTAGAAACCGTGATGATGCAAGAACGAGAGTGTCTTTCATACCAATATATAGCT 465
Qy 481 TTAGAACTTGATTTTCTTAATTCGATGCGCTTTTCGCAATAGAAAACCAAGAAAGTTCCA 540
Db 466 TTAGAACTTGATTTTCTTAATTCGATGCGCTTTTCGCAATAGAAAACCAAGAAAGTTCCA 525
Qy 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGAGATGCTCT 600
Db 526 TTATTTGATGGTATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGAGATGCTCT 585
Qy 601 CTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACAGCTTATTTAGCGGCCAA 660
Db 586 CTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACAGCTTATTTAGCGGCCAA 645
Qy 661 GTGGAACAAACGAGAGATTTCCGACTATTTGCGTAGAATGGTATATAACAGGCTTAAAT 720
Db 646 GTGGAACAAACGAGAGATTTTCCGACTATTTGCGTAGAATGGTATATAACAGGCTTAAAT 705
Qy 721 AGCTTTGAGAGGACAAATGCGCAAGTTGGGTGCGTTATTAATCAATTCGCTAGAGATCTA 780
Db 706 AGCTTTGAGAGGACAAATGCGCAAGTTGGGTGCGTTATTAATCAATTCGCTAGAGATCTA 765
Qy 781 ACGTTAGGGTATTTAGATCTAGTGGCACTATTTCCAAAGCTATGACACTCGCACTTATCCA 840
Db 766 ACGTTAGAGTATTTAGATCTAGTGGCACTATTTCCAAAGCTATGACACTCGCACTTATCCA 825
Qy 841 ATAAATACGAGTGCTCAGTTAAACAGGAAAGTTTATACAGCGAATTTGAGCAACAGGG 900
Db 826 ATAAATACGAGTGCTCAGTTAAACAGGAAAGTTTATACAGCGAATTTGAGCAACAGGG 885
Qy 901 GTAAATATGSCAGTATGAATTTGGTATAATAATGCACTTCGTTTCCGCTATAGAG 960
Db 886 GTAAATATGSCAGTATGAATTTGGTATAATAATGCACTTCGTTTCCGCTATAGAG 945
Qy 961 ACTGCGGTTATCCGAAGCCGCGATCTACTTGAATTTCTAGAAACAACTTATCTATT 1020
Db 946 GCTGCGGCTATCCGAAGCCGCGATCTACTTGAATTTCTAGAAACAACTTATCTATT 1005
Qy 1021 ACTTCATCAGATGGAGTCTACTAGGCATATGACTTACTTGGCGGGGCGACAAATTCAA 1080
Db 1006 GCTTCATCAGATGGAGTATATAGGCATATGACTTATTTGGCGGGGCGACAGATTCAA 1065
Qy 1081 TCTCGGCCAATAGAGGCGGATTAATATCCTCAACGCATGGGTCTACCAATATCTCTATT 1140
Db 1066 TCTCGGCCAATAGAGGCGGATTAATATCCTCAACGCATGGGTCTACCAATATCTCTATT 1125
Qy 1141 AATCCTGTATAGATTTATCATTTCTCTCGAGAGCTATATTTGACCTGAATCATGACGGA 1200
Db 1126 AATCCTGTATAGATTTATCATTTCTCTCGAGAGCTATATTTGACCTGAATCATGACGGA 1185
Qy 1201 GTGCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTTAAT 1260
Db 1186 GTGCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTTAAT 1245
Qy 1261 TTTTAGNACCTCAGATATCTTTTGAAGAGTATCTGCTAACTATAGTCAACCCCTATGAG 1320
Db 1246 TTTTACGAAACCTCAGAATATTTCTGTAGAGGTACCGCTAACTATAGTCAACCCCTATGAG 1305
Qy 1321 TCACCTGGGCTTCAAATTAAGAACTCAGAAACTGAATTAACCAACCAACACAGACGA 1380
Db 1306 TCACCTGGGCTTCAAATTAAGAACTCAGAAACTGAATTAACCAACCAACACAGACGA 1365
Qy 1381 CCAATATATGAATCATATAGTCAATAGTTATCTCATAGAGGCTCATTTTCAAACTCTAGG 1440
Db 1366 CCAATATATGAATCTTACAGTCACAGGTTATCTCATATAGGTATATTTTCAATCCAGG 1425
Qy 1441 GTGCATGTACCAAGTATTTCTTGGAGCGACCGCTAGTGCAGATCGTACAAATACCAATAGT 1500
Db 1426 GTGAATGTACCGGTATATTTCTTGGAGCGCATCTGTAGTGCAGATCGTACGAAATCGATTGGA 1485
Qy 1501 TCAGATAGCATAAACAAATACCAATTTGGTAAATCATTTCAACCTTTAAATTCAGGTACCTCT 1560
Db 1486 CCAATATAGAAATCACCCAAATCCCAATGGTAAAGCATCCGAACTTCTCTCAAGGTACCACT 1545

Qy 1561 GTAGTCAGTGGCCCGAGGATTTTACAGAGGGGATATTAATCCGAACCTAACGTTAATGGTAGT 1620
Db 1546 GTTGTTTAGAGACCAAGATTTTACTGTGGGGATTAATCTTTCGAAGAAACGAATACTGTGGGA 1605
Qy 1621 GTACTAAGTATGGGTCTTAATTTTAATAATACATCAATTAACAGCGGTATCCGCTGAGAGTT 1680
Db 1606 TTTGGACCGATAAGAGTAACCTGTTAACCGGACCAATTAACACAAAGATATCGTATAGATTC 1665
Qy 1681 CGTTATGCTGCTTCTCAAAACATGGTCTCGAGGGTAACGTGTCGAGGGGAGTACTACTTTT 1740
Db 1666 CGTATGCTTCAACTGTAGATTTTGATTTCTTTGTATCAACGTGGAGGTACTACTGTAAAT 1725
Qy 1741 GATCAAGGATTCCTCTAGTACTATGAGTCAAAATGAGTCTTTGACATCTCAATCATTTAGA 1800
Db 1726 AATTTTGTAGTTCTCTAGTACATGAACAGTGGAGCGAATTAATACCGAAATTTTGTG 1785
Qy 1801 TTTGCAAGATTTCTCTGTAGGTATTTAGTGCATCTGGCAGTCAA--ACTGCTCGAATPAAGT 1857
Db 1786 AGACGTGCTTTTACTACACCTTTTACTTTTACACAAATTCAGATATAATTCGAACGTCT 1845
Qy 1858 ATAAGTAAATAATGCAAGGTAGACAAACGTTTCTACTTTGATAAAATTCGAATTCATTCGAAT 1917
Db 1846 ATTCAAGGCTTTAGTGGAAATGCGGAAGTGTATATAGATAAAATTTGAAATTTATTCAGTT 1905
Qy 1918 ACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAAGGCGCAAGAGCGGTGAATGCTCTG 1977
Db 1906 ACTGCAACCTTCGAAGCAGAAATATGATTTAGAAAAGCGCAAGAGCGGTGAATGCTCTG 1965
Qy 1978 TTTACTTAATACCAATCCAAAGAGATTTGAAAACAGATGTGACAGATATCATATTTGATCAA 2037
Db 1966 TTTACTTAATACCAATCCAAAGAGATTTGAAAACAGATGTGACAGATTTATCATATTTGATCAA 2025
Qy 2038 GTATCCAAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTAGATGAAAAGAGAAATTA 2097
Db 2026 GTATCCAAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTGGATGAAAAGAGAAATTA 2085
Qy 2098 CTTGAGAAAGTGAATATTCGAAAACGACTCAGTGTGATGAAAAGAAACTTACTCCAAGATCCA 2157
Db 2086 CTTGAGAAAGTGAATATTCGAAAACGACTCAGTGTGATGAAAAGAAACTTACTCCAAGATCCA 2145
Qy 2158 AACTTCACATCCATCAATAAGCAACAGACTTCTATCTACTTACTTAAGAGCAATTCACATCCAG 2217
Db 2146 AACTTCACATCCATCAATAAGCAACAGACTTCTATCTACTTAAGAGCAATTCACATCCAG 2205
Qy 2218 ACATCTATCCATGAACAACTCTGAACATGGATGGTGGGAAGTGAGAACATTCACATCCAG 2277
Db 2206 ACATCTATCCATGAACAACTCTGAACATGGATGGTGGGAAGTGAGAACATTCACATCCAG 2265
Qy 2278 GAAGGAATGACGTATTTTAAAGAGAAATTCAGTCACTACCGGGGACTTTTAAAGAGTGT 2337
Db 2266 GAAGGAATGACGTATTTTAAAGAGAAATTCAGTCACTACCGGGGACTTTTAAAGAGTGT 2325
Qy 2338 TATCCGAGCTATTTATATCAAAAATAGGAGAGTCGGAATTTAAAGCTTTACTCGCTAC 2397
Db 2326 TATCCGAGCTATTTATATCAAAAATAGGAGAGTCGGAATTTAAAGCTTTACTCGCTAC 2385
Qy 2398 CAATTAAGAGGTATTTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAT 2457
Db 2386 CAATTAAGAGGTATTTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAT 2445
Qy 2458 GCGAAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAAGTTGAA 2517
Db 2446 GCGAAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAAGTTGAA 2505
Qy 2518 AGCCCAATTCGAAGGTGCGGAGAACCGAATTCGATGCGCACCAATTTTGAATGGAATCCT 2577
Db 2506 AGCCCAATTCGAAGGTGCGGAGAACCGAATTCGATGCGCACCAATTTTGAATGGAATCCT 2565
Qy 2578 GATCTAGATTTCTCTCAGAGTGGAGAAAATGTCGCGCATCTTCCCATCATTTCTCT 2637
Db 2566 GATCTAGATTTCTCTCAGAGTGGAGAAAATGTCGCGCATCTTCCCATCATTTCTCT 2625

```
Qy 2638 TTGGATATTGATATTGGATGCAACAGCTTGCAATGAGAACTAGGCGTGTGGTGTATTC 2697
Db |||||
Qy 2626 TTGGATATTGATTTGGATGCAACAGCTTGCAATGAGAACTAGGCGTGTGGTGTATTC 2685
Db |||||
Qy 2698 TAGATTAGACGACGAGGTCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA 2757
Db |||||
Qy 2686 AAGATTAGACGACGAGGTCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA 2745
Db |||||
Qy 2758 CCAATTATTAGGAGGACACTGCTCGTGTGAAGAGACGAGAGAAAAAATGGAGACAAA 2817
Db |||||
Qy 2746 CCAATTATTAGGAGGACACTGCTCGTGTGAAGAGGACGAGAGAAAAAATGGAGACAAA 2805
Db |||||
Qy 2818 CGTGAATACTCAATTTGAAACAAAACGAGTATATACAGAGGCAAAAGAGCTGTGGAT 2877
Db |||||
Qy 2806 CGTGAATACTCAATTTGAAACAAAACGAGTATATACAGAGGCAAAAGAGCTGTGGAT 2865
Db |||||
Qy 2878 GCTTTATTGTAGATTCTCAATATATAGATTACAGCGGATACAAACATTTGGCATGATT 2937
Db |||||
Qy 2866 GCTTTATTGTAGATTCTCAATATATAGATTACAGCGGATACAAACATTTGGCATGATT 2925
Db |||||
Qy 2938 CATGCGGACAGATAAACTTGTTTCATCGAATTCGAGAGGCTTATCTGTCAGAATTTATCTGTT 2997
Db |||||
Qy 2926 CATGCGGACAGATAAACTTGTTTCATCGAATTCGAGAGGCTTATCTGTCAGAATTTATCTGTT 2985
Db |||||
Qy 2998 ATCCCGGCTGTAATTCGGAATAATTTTGAAGAAATTTAGAGGTCGCATTATCACTGCAATC 3057
Db |||||
Qy 2986 ATCCCGGCTGTAATTCGGAATAATTTTGAAGAAATTTAGAGGTCGCATTATCACTGCAATC 3045
Db |||||
Qy 3058 TCCCTATACGATCGAGAAATGTCGTTAAATAATGTTGATTTTAAATGGAATTTAGCATGC 3117
Db |||||
Qy 3046 TCCCTATACGATCGAGAAATGTCGTTAAATAATGTTGATTTTAAATGGAATTTAGCATGT 3105
Db |||||
Qy 3118 TGGAAATGTAAGGCGCATGTAGATGTCAACAGAGCCATCAACGCTTCTGCTCTGTTTATC 3177
Db |||||
Qy 3106 TGGAAATGTAAGGCGCATGTAGATGTCAACAGAGCCATCAACGCTTCTGCTCTGTTTATC 3165
Db |||||
Qy 3178 CCAGAAATGGAAGAGAGAGTGTCAACAGAGTTCGCGTCTGTCCGGGCGTGGCTATATC 3237
Db |||||
Qy 3166 CCAGAAATGGAAGAGAGAGTGTCAACAGAGTTCGCGTCTGTCCGGGCGTGGCTATATC 3225
Db |||||
Qy 3238 CTCGCTGTCAAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAATC 3297
Db |||||
Qy 3226 CTCGCTGTCAAGCGGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAATC 3285
Db |||||
Qy 3298 GAGAAATACAGAGCAACTAAATTTTAAATACTGTGAAGAGAGGAGTGTATCCCAAG 3357
Db |||||
Qy 3286 GAGAAATACAGAGCAACTAAATTTTAAATACTGTGAAGAGAGGAGTGTATCCCAAG 3345
Db |||||
Qy 3358 GATACAGGAACGTGTAATGATTATATCTGCACACCAAGGTACAGC-----AGTA 3405
Db |||||
Qy 3346 GATACAGGAACGTGTAATGATTATATCTGCACACCAAGGTACAGCTGGATGCGCAGATGCA 3405
Db |||||
Qy 3406 TGTAATTTCCGTAATGCTGGATATGAGATGCAATGAAAGTGTGATCTACAGCATCTGTT 3465
Db |||||
Qy 3406 TGTAATTTCCGTAATGCTGGATATGAGATGCAATGAAAGTGTGATCTACAGCATCTGTT 3465
Db |||||
Qy 3466 AATTACAAACCGACTTATGAAGAGAAACGATATACAGATGTACGAAGAGATAATCATTTG 3525
Db |||||
Qy 3466 AATTACAAACCGACTTATGAAGAGAAACGATATACAGATGTACGAAGAGATAATCATTTG 3525
Db |||||
Qy 3526 GAATATGACAGAGGATGTCAAATTTATCCACCAAGTACCAGCTGGTTATGTGACAAAAGAA 3585
Db |||||
Qy 3586 TTAGAATACTTCCGAAACCGGATAGGTATGGAATTTGAGATTTGAGAAACCGAAGGGAAG 3645
Db |||||
Qy 3586 TTAGAATACTTCCGAAACCGGATAGGTATGGAATTTGAGATTTGAGAAACCGAAGGGAAG 3645
Db |||||
Qy 3646 TTTATTGTAGACAGCGTGAATTTACTCTCTTATGAGGAATAG 3687
Db |||||
Qy 3646 TTTATTGTAGACAGCGTGAATTTACTCTCTTATGAGGAATAG 3687
Db |||||
```

RESULT 5

```
US-09-826-660-22
; Sequence 22, Application US/09826660
; Patent No. US20010026940A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XCD1
; CURRENT APPLICATION NUMBER: US/09/826,660
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 3558
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-826-660-22
```

```
Query Match 42.9%; Score 1581.6; DB 9; Length 3558;
Best Local Similarity 66.5%; Pred. No. 0;
Matches 2458; Conservative 0; Mismatches 1084; Indels 156; Gaps 7;

Qy 2 TGACTTCAATAGGAAAATGAGATGAAATTAATAATGCTTTATCGATTCACGCTGAT 61
Db 2 TGACTTCAATAGGAAAATGAGATGAAATTAATAATGCTTTATCGATTCACGCTGAT 61

Qy 62 CGAATCAATTCACACAAATGGATCTATCACAGATGCTCGTATTGAGGATTCCTTTGTGTA 121
Db 62 CTAACCAATTCGCTCAGATGAACTTTCTACTGATCTAGATCGAGGATTCCTTTGTGTA 121

Qy 122 TAGCCGAGGGAATAATATCAATCCACTTTGTTAGCGATCAACAGTCCAAACGGGTATTA 181
Db 122 TTGCTGAGGGAACAACAAATTCGATTCCTTTCTGCTTCTACTGTTCAAACTGGAATCA 181

Qy 182 ACATAGCTGTAGAAATGATAGTGTATAGCGGTACCGTTTCTGCTGGACAAATAGTAGTT 241
Db 182 ACATTTGCTGGAAGAAATCCTTTGGAGTTCCTTGGAGTTCATTCGCTGGACAGATTCCTTT 241

Qy 242 TTTATAGTTTCTTTGTTGGTGAATTTATGGCCCGCCGCGCAGAGATCAGTGGGAAAATTTTCC 301
Db 242 TCTACTCTTTCTTTGTTGGAGAGCTTTGGCTTAGGGGAAGAGATCCTTTGGAGATCTTCC 301

Qy 302 TAGAATATGTCGAACAACTTATATAATCAACAAATPAAACAGAAAAATGTAGGAATACGGCAC 361
Db 302 TTGAGCATGTTTGAGCAGTTGATTCGTCACAAAGTTACTTGAGAACACTAGAGATACTGCTC 361

Qy 362 TTGCTCGATTAACAGGTTTATAGGAGATTCCTTTAGAGCTATCAACAGTCACTTCGAAGATT 421
Db 362 TTGCTAGACTTCAAGGACTTTGGAACCTTTTCAGAGCTTACCAACAAATCTCTTGAGGATT 421

Qy 422 GGCTAGAAAAACCGTGTAGTATGCAAGAACGAGAAAGTGTTCCTTTATACCAATATATAGCCT 481
Db 422 GGCTTAGAAGACAGAGATGATGCTAGAACTAGATCTGTGTTGTACTACTCAGTACATTTGCTC 481

Qy 482 TAGAATCTGATTTTCTTAATGCGATGCGCTTTTTCGCAATTTAGAAAAACGAGAAAGTTCCAT 541
Db 482 TTGAGCTTGACTTCTTGAACGCTATGCGCATTTGTTTCGCTATCAGAAAAACGAGAGTTCCAC 541

Qy 542 TATTTAATGTTATATGCTCAAGCTGCAAAATTTTACACCTTATTTATTTCAGAGATGCTCTC 601
Db 542 TTCTCATGGTGTACGCTCAAGCTGCAAACTTCACTCTTCTTCTTCTTAGAGATGCTTAGCT 601

Qy 602 TTTTGTGTAGTAATTTGGGCTTACATCGCAGGAAATTTCAACGTTTATTATTAGCGCCAAG 661
```

Db 602 TGTTCGGATCTGAGTTCGGACTTACTCTCAAGAGATTCAAGATATCTACGAGAGCAAG 661
Qy 662 TGGAAACAAACGAGAGATTATTCGAGTATTCGCTAGAAATGGTATAATACAGGTCTAAATA 721
Db 662 TTGAGAAGACTAGAGAGTACTCTGACTACTCGCTAGGTGGTACAAACACTGGATTGAACA 721
Qy 722 GCTTCGAGAGGACAAATCCGCAAGTTCGGTGGCTTATATCAATTCGCTAGAGATCTAA 781
Db 722 ACCTTAGAGGAACTAACGCTGAGTCTGGCTTAGATACACCACTGAGATTCAAGAGAGATCTTA 781
Qy 782 CGTTAGGGGTATTTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTTATCCAA 841
Db 782 CTCCTGGAGTCTTGATCTTGCTTGCTTGCTTCCCATCTTAGAGATCTAGAGTGTACCCCTA 841
Qy 842 TAAATACAGATGCTCAGTTAAACAGGGGAAGTTTATACAGCGCAATTTGGAGCAACAGGG 901
Db 842 TGAACACTTCTGCTCAACTTACTAGAGAGATCTACACTGATCCAAATCGGAAGAACTAACG 901
Qy 902 TA-----AATATGCGAAGTATGAATGGTATATATATATGATGACACTCGCACTTTATCCAA 955
Db 902 CTCATCTGGATTGCTTCTACTAACTGGTTCACAAACAGCGTCCATCTTTCTCTGCTA 961
Qy 956 TAGAGATCGCGGTATTCGAAAGCCGCACTACTTTGATTTTCTAGAACAACTTACAAATTT 1015
Db 962 TCGAGGTGCACTGATCAGACCACCACTCTTCTTGACTTCCCGAGACAACTTACTATCT 1021
Qy 1016 TTAGCACTTCATCAGATGGAGTCTACTAGGCATATGACTTACTGCGGGGGGACACAA 1075
Db 1022 TCTCTGTCTTCTTAGATGGTCTAACCACTCAGTACATGAATCTACTGGGTTGGACATAGAC 1081
Qy 1076 TTCAATCTCGGCCAATAGGAGCGGATTAATACTCAACGCAATGGGTCTACCAATACTT 1135
Db 1082 TTGAGTCTAGAACTATCAGAGGATCTCTTCTACTTCTACTCTCATGGAAACACTTAACACTT 1141
Qy 1136 CTATTATCTGTGAAGATTATCATCTTCTCTGAGACGCTATATTTGGAATCATATATG 1195
Db 1142 CTATCAACCCAGTTACTCTTCAGTTCTACTCTAGAGATGGTACAGAACTGAGTCTTTCG 1201
Qy 1196 CAGGAGTGTCTTATGGGAATTTACCTGCACTTATGATGTCCTTACTGCTTACTGAT 1255
Db 1202 CTGGAATCAACAT-----TCTTCTACTACTCAAGTGAACGGAGTTCCTTGGGCTAGAT 1255
Qy 1256 TTAATTTTAGAACCCCTCAGAATACTTTTGAAGAGTACTGCTAATATAGTCAACCCCT 1315
Db 1256 TCACTGGAGAAACCCATTGAACCTCTT--AGAGGTTCTTGTGTACACCACTTGGAT 1312
Qy 1316 ATGAGTCACTGGGCTTCAATTAAGATTCAAGAACTGAATTAACCAACGAAACAAACAG 1375
Db 1313 AACTGGAGTTGGTACCCAGTTGTTCGATTTCTGAGACTGAGCTTCCACCAGAGACTACTG 1372
Qy 1376 AACGACCAAAATTATGAATCATATAGTCAATAGTTTATCTACATAGGGCTCATTTTCAAAAT 1435
Db 1373 AGAGACCAAACTACGAGTCTTACTCTCATAGACTTCTTAACATCTGTTGATCTCTGGAA 1432
Qy 1436 CTAGGGTGCATGATACAGTATATCTTTGGACGACCGTAGTCAGATCGTACAAATACCA 1495
Db 1433 ACACCTCTTAGAGTCCAGTGTACTCTTGGACTCATAGATCTGCTGATAGAACTTAACACCA 1492
Qy 1496 TTAGTTCAGATAGCATAAACAAATACCATTTGTTAAATCATTTCAACCTTAATTCAGGTA 1555
Db 1493 TCTCTTCTGATTTCTATCACTCAGATTCCACTTGTGAAGTCTTTCAACTTGAACCTCTGGAA 1552
Qy 1556 CTTCTGTAGTCACTGGCCCAAGGATTTACAGGAGGGGATATAATCCGAACCTAACGTTAATG 1615
Db 1553 CTTCTGTGTTCTTGGACCAAGGATTCACCTGGAGGAGACATCATCAAACTAACGTGAACG 1612
Qy 1616 GTAGTGTACTAAGTATGGGCTTAATTTAATTAATCATCATTAACGCGGTATCGCGTA 1675
Db 1613 GATCTGTCTTCTATGGGAATGAATCTTCAACAAACACTTCTCTTCAAGATACAGAGTTA 1672
Qy 1676 GAGTTCGTTATGCTGCTTCTCAAACTATGCTGAGGGTAACTGTCGGAGGAGTACTTA 1735

Db 1673 GAGTTAGATACGCTGCTTCTCAAACTATGTTCTTAGAGTTACTGTTGGAGGATCTACTA 1732
Qy 1736 CTTTGTATCAAGGATTCCTCTAGTACTATGAGTCAAAATGAGTCTTTGACATCTCAATCAT 1795
Db 1733 CTTTCGATCAAGGATTCCTCTAGTACTATGTTCTGCTAACGAGTCTCTTACTTCTCAATCTT 1792
Qy 1796 TTAGATTTGAGAAATTTCTCTGTAGTATTTAGTGCATCTGCGAGTCAAACTGCTGGAATAA 1855
Db 1793 TCAGATTCGCTGAGTTCCAGTTTGGAAATCTCTGCTTCTGATCTCAAACTGCTGGAATCT 1852
Qy 1856 GTATATAGTATATGCAAGTACAGTACAAACGTTTTCACCTTTGATATAAATGAAATTCATTCCAA 1915
Db 1853 CTATCTCTAACCAACGCTGGAAGACAAACTTTCACCTTCGACAAAGATTGAGTTTCAATTCCAA 1912
Qy 1916 TTACTGCAACCTTTCGAAGCAGAAATACGATTTAGAAAGGCGCAAGAGGCGGTGAATGCTC 1975
Db 1913 TCACTGCTACTCTCGAGGCAGAGTCTGACTTGGAAAGAGCACAGAAGGCGGTGAATGCTC 1972
Qy 1976 TGTTTACTTAATAGAAATCCAAAGAGATTGAAACAGATGTGACAGATTAATCATATTTGATC 2035
Db 1973 TGTTCATCTTCGTCATCATGATTTGGGCTCAAGACAGATGTGACTGACTATCATCATGATC 2032
Qy 2036 AAGTATCCAAATTTAGTGGGCTGTTTATCGATGAATTTCTGCTTAGATGAAGAGAGAAAT 2095
Db 2033 GCGTTTCCAACTTGTGAGTGCCTCTCTGATGAGTTCGTGTTGGATGAAGAAAGAGAT 2092
Qy 2096 TACTTGAGAAAGTGAATATGCGAAACGACTCAGTGAATGAAGAAACCTTACTTCCAGATC 2155
Db 2093 TGTCCGAGAAGGTCAAAACATGCTAAGCGACTTAGTGTGATGAGCGAACTTGTCTTCAAGATC 2152
Qy 2156 CAAACTTCACATCCATCAATTAAGCAACCAAGACTTTCATATCTACTAATGAGCAATCGAAT 2215
Db 2153 CCAACTTTCGCGGGATCAACAGGCAACTA----- 2181
Qy 2216 TCACATCTATCCATGAACAACTCTGAACATGGATGGTGGGAAAGTGAGAACTTACAATCC 2275
Db 2182 -----GATCGTGGATGGAGGGAAGTACGGACATCAACATTC 2218
Qy 2276 AGGAAGAAATCAAGTATTTAAAGAAATTAAGTCACTACCGGGGACTTTTAATGAGT 2335
Db 2219 AAGGAGGTGATGATGTTTCAAGAGGAATATGTTACGCTCTTGGGTACTCTTGTATGAT 2278
Qy 2336 GTTATCCGACGCTATTTATATCAAAAAATAGGAGAGTTCGGAATTTAAAGCTTATATCTCGCT 2395
Db 2279 GCTATCCAACTACCTGTACCAAGAGATAGATGAATCGAACTCAAGGCTTACAAAGAT 2338
Qy 2396 ACCAATTAAGGGGTATATTTGAAGATAGTCAAGATTTTAGAGATATATTTGATTCGTTATA 2455
Db 2339 ACCAGTTGAGAGGTTACATCGAGGACAGTCAAGACCTTGAGATCTACCTCATCAGATACA 2398
Qy 2456 ATGCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGTATGGCGGCTTTTCAAGTTG 2515
Db 2399 AGCCAAACATAGACAGTCAATGTGCTGGGACGGGTTCACTCTGGCCACTTTTCAGGCC 2458
Qy 2516 AAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCACCAACATTTTGAATGGAATC 2575
Db 2459 CAAGTCCCATCG----- 2471
Qy 2576 CTGATCTAGATTTGTTCTCTGAGAGATGGAGAAAATGTGCGCATCATTTCCCATCTTTCT 2635
Db 2472 -----CAAGTGTGCCCATCACTCACACCACTTCT 2500
Qy 2636 CTTTGGATTTGATATTTGGATGCAAGACTTTCATGAGAAATCTAGGCGGTGGGGTGTAT 2695
Db 2501 CTTTGGACATAGACGTTGGCTGTACCGACCTGAACGAAGACCTCGGTGTGGGTGATCT 2560
Qy 2696 TCAAGATTAAGACGAGGAGGTCATGCAAGACTAGGGAATCTCGAAATTTTATTTAGAGAGA 2755
Db 2561 TCAAGATCAAGACTCAAGATGCCCATGCGAGCTAGGCAATCTCGAGTTCTTAGAAGAGA 2620
Qy 2756 AACCAATTAAGGAGAGCACTGCTGCTGTTGAGAGAGCAGAGAAAATTTGGAGAGACA 2815
Db 2621 AACCACTGTTGGAGAAAGCCCTCGCTAGAGTGAAGAGGGCTGAGAAAGAGTGGAGGGACA 2680

Qy	259	GGTGAATTTATGGCCCGCGGCGAGAGATCAGTGGGAAATTTTCTAGAAACATGTCTGCAACAA	318
Db	181	GGCGAGCTGTGGCCCGCGGCGCGACAGTGGGAGATCTTCTCGAGACACGTGGAGCAG	240
Qy	319	CTTATAAATCAACAATAACAGMAAATCTAGGATACGGCACTTGTCTCGATTACAAGT	378
Db	241	CTGATCAACACGACGATCACCGAGAACCGCCGCAACCGCCCTGGCCGCGCTCGAGGC	300
Qy	379	TTAGGAGATTCTTTTAGAGCCTTATCAACAGTCACCTTGAAGATTGGCTAGAAAACCGTGAT	438
Db	301	CTGGCGCAGAGTTTCCGCGCCTACACGACAGACCTGGAGGACTGGCTGGAGAACCGCGAC	360
Qy	439	GATCAAGAAACGAGAGTGTCTTTATACCCNAATATAGCCTTAGAACTTGAATTTCTT	498
Db	361	GACGCCCGCAACCGCAGCGTGTGTACACCCAGTACATCGCCCTCGAGCTCGACTTCCTG	420
Qy	499	AATCGATCGCGCTTTTTCGCAATTAGMAACCAAGATTCATATTAAATGATATATGCT	558
Db	421	AACGCCATGCCCTGTTTCGCCATCCGCAACACGAGGTTGCCCTGCTGATGGTGTACGCC	480
Qy	559	CAAGCTGCMAATTTACACCTATTATTATGAGAGATGCCTCTCTTTTGGTAGTGAAATTT	618
Db	481	CAGCGCGCAACCTTGCACTCTGCTGTGTGCGCAGCACGCCCTGTTCTGGCAGCGAGTTC	540
Qy	619	GGGCTTACATCGCAGGAAATTCACAGTTATTATGAGCGCCAAAGTGGAAACAAACGAGAGAT	678
Db	541	GGCCTGACACGACGAGAGATCCAGGCTACTACGAGCGCAGGTGAGCGCACCCGCGAC	600
Qy	679	TATTCGACTATTCGGTAGAATGTATAATACAGGTCTAAATAGCTTGAGAGGACAAAT	738
Db	601	TACAGCACTACTCGTGGAGTGTACAACACCGGCTGTAAACGCTGCGCGCACCAAC	660
Qy	739	GCCCAAGTTGGGTCGTTAATCAATCCGTAGAGATCTAAGTTAGGGTATTAGAT	798
Db	661	GCCGCAAGTCGGTGCCTACACGATTTCGCGCGCACTGACCTTGGGGCTGCTGGAC	720
Qy	799	CTAGTGCACTATTCCTCAAGCTATGACATCGCACTTATCCAAATAAATACAGATGCTCAG	858
Db	721	CTGTGTGCCCTGTTTCCCGAGCTACGACACCGCGACCTACCCCATCAACACCGCGCCAG	780
Qy	859	TTACAAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGGTAAATATGGCAAGTATG	918
Db	781	CTGACCCGAGGTGTACCGACGCACTCGCGCCACCGCGGTGAACATGCGCAGCATG	840
Qy	919	AATTTGATATAATAATATGCACTTCGTTTTTCGCTATAGAGACTGGGGTTATCCGAAGC	978
Db	841	AACTGTGTACAACAACACGCCCCAGCTTCAGCGCCATCGAGGCGCGCGCATCCGACGC	900
Qy	979	CCGATCTACTTGATTTTCTAGAACAACTTACAAATTTTATAGCACTTCATCAGATGAGT	1038
Db	901	CCCCACTGTGTGACTTCTGTGAGACGTGACCAATCTTCAGCGCCAGCAGCCGCTGGAGC	960
Qy	1039	GCTACTAGGCATATGACTTACTGGCGGGGGCACAAATTCAAATCTCGGCCAAATAGGAGGC	1098
Db	961	AACACCGCCACATGACCTACTTGGCGGGCCACACCATCAGAGCCGCCCATCGCGCGC	1020
Qy	1099	GGATTAATACCTCAAGCATGGGTGTACCAATATCTTCTATTAATCTGTAAAGATTATCA	1158
Db	1021	GGCCTGAACACACGACCCACCGGGCCACCAACACACAGCATCAACCCCGTGAACCCCTGCGC	1080
Qy	1159	TTCTTCTCTCGAGACGTATATTGACATGAAATCATATGACAGAGTCTTCTATGGGAATTT	1218
Db	1081	TTGCCAGCCGACGCTGTACCGCACCGAGCTATCGCGGCGTGTCTGTGGGGGATC	1140
Qy	1219	TACTTTGAACCTATTCATGGTGCCTTACTGTGTAGATTAAATTTTAGGAACCCCTCAGAAT	1278
Db	1141	TACTGTGAGCCCATCCACGGCGTGCACCGTGGCTTCAACTTCACCAACCCCCAGAAC	1200
Qy	1279	ACTTTTGAAGAGTACTGTCTACTATAGTCAACCCCTATGATCACTGGGCTTCAATTA	1338
Db	1201	ATCAGGACCGCGCACCGCAACTACGCGAGCCCTACGAGAGCCCGCGCTGCGAGCTG	1260
Qy	1339	AAGATTTCAGAAACTGAATTTACCACCAAGAAACACAGAACCGACCAATTTATGAATCATAT	1398

[illegible]

Db 2341 GAGGACAGCAGGACCTGGAGATCTACCTGATCCGTACAACGCCAAGCAGCAGACCCCTG 2400
Qy 2476 GATGTTTCAGGTACCGAGTCCGTATGCGCGCTTTTCAGTTTGAAGCCCAATCGGAGGTGC 2535
Db 2401 GACGTGCCGCGCACCGAGGCTGTGGCCCTTGAGCGTGGAGAGCCCATCGCGCGCTGC 2460
Qy 2536 GGAGAACCGAATPCGATGCGCACACATTTTGAATGGAATCCCTGATCTAGATTGTTCTTCGC 2595
Db 2461 GCGGAGCCCAACCGCTGCGCCGCCACTTCGAGTGGAAACCCCGACCTGGAAGCTGCAGCTGC 2520
Qy 2596 AGAGATGGAAGAAATGCGCATCATTTCCCATCATTTCTCTTTGGATTTGATATTGGA 2655
Db 2521 GCGGACGCGGAGAGTGGCCGCCACACACAGCCACCACTTCAGCCTGGACATCGACGTGGGC 2580
Qy 2656 TGCACAGACTTGCATGAGAACTACGCGTGGGTGGTATTCAAGATTAAAGACGAGAA 2715
Db 2581 TGCACGACCTGACAGAACTCGGGCGTGGGTGGTTCAGATCAGAACCCGAGAG 2640
Qy 2716 GGTATGCAAGACTAGGGAATCTGGAATTTATGGAAGAAACCAATATTATAGGAGAAGCA 2775
Db 2641 GGCACGCGCGCTGGGCAACTGGAGTTTCATCGAGGAGAAAGCCCTGCTGGCGAGGCC 2700
Qy 2776 CTGCTCTGTGTAAGAGCAGAGAAATAATGGAGAGCAAAACGTGGAATACTAATTTG 2835
Db 2701 CTGAGCCCGTGAAGCGCGCGAGAAAGTGGCGCGAACAGCGCGAGAAGCTGCAGCTG 2760
Qy 2836 GAAACAAACAGAGTATATACAGAGCAAAAGAGCTGTGGATGCTTTATTTGTAGATTCT 2895
Db 2761 GAGACCAAGCGGTGTACACGAGGCCAAGAGGCCGTGGAGCCCTGTTTGTGGACAGC 2820
Qy 2896 CAATATAATAGATTACAAGCGGATACAACATTTGGCATGATTCATCGCGCAGATAAACTT 2955
Db 2821 CAGTACGACCGCTCGAGCGCGACACCAACATCGGCATGATCCAGCGCGCGACAAGCTG 2880
Qy 2956 GTTCATCGAATTCGAGAGCTTATCTGTCAAGATTATCTGTTATCCCGGGTGTAAATCGG 3015
Db 2881 GTGCACCGCATCCGCGCGCTTACCTGAGCGAGCTGCCCGGTGATCCCGCGGTGAACGCC 2940
Qy 3016 GAAATTTTGAAGAAATTAGAGTTCGCAATTCATCTGCAATCTCCCTATACGATCGGAGA 3075
Db 2941 GAGATCTTCGAGGAGCTGGAGGCCACATCATCACGCCATCAGCTGTACAGCCCGC 3000
Qy 3076 AATGCTGTTAAATAATGTTGATTTTAAATGGAATPAGATGCTGGAATGTAAAGGGCAT 3135
Db 3001 AACGTGTGAAGAACCGGACTTCAACAAAGCGCTGACCTGCTGGAAGCTGAAGGGCCAC 3060
Qy 3136 GTAGATGTACAACAGAGCCATCACCGTCTGCTCTGTTATCCAGAAATGGGAAGCAGAA 3195
Db 3061 GTGGACGTGCAGCAGAGCCACCCCGCAGCAGCACTGGTGTATCCCGAGTGGGAGCCGAG 3120
Qy 3196 GTGTCAAGCAGTTTCGCTGTCGCGGGGTGGCTATATCTCCGTGTCACAGCGTAC 3255
Db 3121 GTGAGCCAGCGCTGCGGTGTGCCCCGGCTGCGGCTACATCTCTCGCGGTACCGCCTAC 3180
Qy 3256 AAAGAGGATATGGAGAGGGTGTGTAAACGATCCATGAAATCGAGAAACAATACAGAGAA 3315
Db 3181 AAGGAGGGCTACGGCGAGGGCTGCTGACCATCCACAGAGATCGAGAAACAACAGCAGAG 3240
Qy 3316 CTAAATTTAAATCTGTAAGAGAGAGAGTGTATCCAGCGGATACAGGAACGTGTAT 3375
Db 3241 CTGAAGTTCAAGAAACCGGAGGAGGAGGTGTATCCCGCAGCACCGGACCTGCAAC 3300
Qy 3376 GATTATCTGCACACCAAGGTACAGC-----AGTATGTAATTCGCGTAATGCT 3423
Db 3301 GACTACACCGCCACAGGGACCCCGCTGCGCGGACCGCTGCAACAGCGCGCAACGCC 3360
Qy 3424 GGATATGAGGATGCAATATCAAGTTGATATACAGCATCTGTTAATTAACAAACCGACTTAT 3483
Db 3361 GGTACAGAGGAGCGCTACAGGTGGACACACCGCGAGCGTGAATCAAGCCCACTAC 3420
Qy 3484 GAAGAGAAACGTATACAGATGTACGAAGAGATATCATTTGTGATATGACAGAGGGTAT 3543
Db 3421 GAGGAGGAGACCTTACACCGAGTGGCGCGCAGCAACCACTGCGAGTACGACCGCGGCTAC 3480

Qy 3544 GTGAATTTATCCACCACTACCACTGGTTTATATGACAAAGAAATTAGAAATCTTCCAGAA 3603
Db 3481 GTGAATTTATCCACCACTGGTTTATATGACAAAGAAATTAGAAATCTTCCAGAA 3540
Qy 3604 ACCGATAGGTATGATTTGAGATTGGAGAAACGGAAGGGAAGTTTATTTGTAGACAGCGTG 3663
Db 3541 ACCGACACCGTGTGGATCGAGATCGGCGAGACCGGAGGCAAGTTTCATCTGTGACAGCGTG 3600
Qy 3664 GAAATTTATCTTATGAGGAGGAGTATG 3687
Db 3601 GAGTGTCTGTATGAGGAGGAGTATG 3624

RESULT 7

US-09-826-660-5
; Sequence 5, Application US/09826660
; Patent No. US20010026940A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; CURRENT APPLICATION NUMBER: US/09/826,660
; FILE REFERENCE: MA-714XC2D1
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3522
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-826-660-5

Query Match 38.2%; Score 1408.4; DB 9; Length 3522;

Best Local Similarity 65.7%; Pred. No. 0;

Matches 2243; Conservative 0; Mismatches 1061; Indels 108; Gaps 9;

Qy 285 TCAGTGGGAAATTTCTTAGAACATGTCGAACAATTTATTAATCAACAAATAACAGAA 344
Db 207 TGATTTGAGCTTATTTCTTTTACAGATTGAACAATTTGATGAGCAAGAAAGAAACATT 266
Qy 345 TGCTAGGAATACGCGACTTGTCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTATCA 404
Db 267 GGAAGGAAACGGGCAATTTACTACATTACGAGGCTTAGCAGATAGTAAATTTATAT 326
Qy 405 ACAGTCACTTGAAATTTGGCTAGAAACCGGTGATGATGCAAGAAAGAGAAAGTGTCTTTA 464
Db 327 TGAAGACCTAAGAGAGTGGGAAGCAATCTCTAATATGTCACAAATTAAGGGAAGATGTCG 386
Qy 465 TACCAATATATATAGCTTTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAG 524
Db 387 TATTCGATTTGCTTAATACAGACGAGCTTTTAATACAGCAATAAATAATTTTACACTTAC 446
Qy 525 AAACCAAGAAAGTTCCTATTTAATGTTATATGCTCAAGCTGCAAAATTTTACACCTATTATT 584
Db 447 AAGTTTTGAAATCCCTTTTATCGGCTATGTTCAAGCGGCGAATTTACATTATCACT 506
Qy 585 ATTGAGAGATGCTCTCTTTTGTGATGAAATTTTGGGCTTACATCGCAGGAAATTTCAACG 644
Db 507 ATTAAGAGACGCTGTATCTGTTTGGGCGAGGTTGGGACTTGGATATAGCTACTGTTAATAA 566
Qy 645 TTATTATGAGCGCCCAAGTGGGAACAACGAGAGATTTATCCGACTATTTCGCTAGAGTATGTA 704
Db 567 TCATTATAATAGATTAATAAATCTTTATTCATAGATATACGAAACATTTGTTGGACACATA 626
Qy 705 TAATACAGGTCTAAATAGCTTTGAGAGGGAACAAATGCCGCAAGTTGGGTGCGTTAATATCA 764

	Db	627	CAATCAAGGANTTAAAGAACTTAGAGGTACTAATACTCGACAATGGCAGATTCAATCA	686
	Qy	765	ATTCCGTAGAGATCTTAACGTTAGGGGTATTAGATCTAGTGGCACATATTCCTCCAGGTATGA	824
	Db	687	GTTTAGGAGAGATTAAACACTTACTGTATTAGNATCGTTGCTCTTTTTCCGAACCTAGA	746
	Qy	825	CACCTCCGACTTATCCAATAAATACGAGTGCTCAGTTAACAGAGGAAGTTTATACAGAGC	884
	Db	747	TGTTAGAACATATCCAATTTCAAACGTCATCCCAATTAAACAAGGAAAATTTATACAAGTTC	806
	Qy	885	AATTGGAGCAACAGGGGTAAAAATATGCAAGTATGAATTGGTATATAATATATGCACTTC	944
	Db	807	AGTAATTGAGGATTCTCGAGTTTCTGCT-----TAATATACCTTAA	845
	Qy	945	GTTTTCCGCTATFAGAGACTCGCGTTATCCGAAGCCCGCATCTACTTGATTTTCTTAGAACA	1004
	Db	846	TGGTTTTAATAGGGGGAAATTTGGAGTTAGACCGCCCCCATCTTATGGACTTTATGAA---	902
	Qy	1005	ACTTACAATTTTTAGCACCTTCATACGATGGAGTGCTACTAGGCCATATGACTTATGGCG	1064
	Db	903	-----TTCTTTGTTGTAACTCGACGAGACTGTTAGNAGTCAAACTGTGTGGGAGG	953
	Qy	1065	GGGGCACAAATTCATCTCGGCCAATAGGAGCGGATTAATACTCTCAACGCATGGGTC	1124
	Db	954	ACACTTAGTTAGTTCACGAATAACGGCTGGTAAACCGTATAAATTTCCCTAGTTAGGGGT	1013
	Qy	1125	TACCRAATCTTCTATTATCTGTGAAGATTAATCATTTCTTCTCTCGAGACGPATATTTGAC	1184
	Db	1014	CTTCAATCTGGTGGCGCATTTGGATTTCGAGATGAGGATCCAGTCTCTTTTATTCGAC	1073
	Qy	1185	TGAATCATATGACGAGTGCTTCTATGGGGAATTTACCTTCAACCTATTTCATGTGTCCC	1244
	Db	1074	ATTATCAGATCCTGTTTTGTCGGAGGAGAT-----TTGGGAATCTCTAATATGACT	1127
	Qy	1245	TACTGTTAGATTAATTTTTAGAAACCTTCAGAACTACTTTTGAAAGAGGTACTGCTAACTA	1304
	Db	1128	GGGGCTTAGGGGAGTAGCAITTTCAACAACTGGTAGCAACCAACCCGAACTATTAGAAA	1187
	Qy	1305	TAGTCAACCCCTATGAGTCACCTGGGCTTCAATTTAAAGATTCAGAAACTGNAATTACCAC	1364
	Db	1188	TAGTGGGACCATAGATTTCTCTAGATGAATCCCACTCAGGATTAATAGTGGGGCACC--TT	1246
	Qy	1365	AGAAACAACAGAACGACCAAAATATGAATCATATAGTCATAGGTTATCTCACATAGGGCT	1424
	Db	1247	GGAAATGATATAGTCATGTATTAATATCAATGTTACATTTGTACGATGCCAGGTGAGATTT	1306
	Qy	1425	CATTTCAAACTTAGGTGTCATGACGATATATTTCTTGACGACACCGTAGTCAGATCG	1484
	Db	1307	CAGGAAGTGATTCATGAG--AGCTCCATGTTTTCTTGACGACACCGTAGTGCAACCCC	1364
	Qy	1485	TACAAATACCAATAGTCAGATAGCATAAACAATAACCATACCTTGGTAAAACTCATTTCAACCT	1544
	Db	1365	TACAAATACCAATGATCCGAGAGGATTAATCTCAAAATACCATGGTAAAGACCATACACT	1424
	Qy	1545	TAATTCAGGTACCTCTGTCAGTGGCCGACGATTTACAGGAGGGGATATAATCCGAAC	1604
	Db	1425	TCAGTCAGGTACTACTGTTGAAGAGGGCCCGGTTTACGSGAGGAGATATTTCTCGACG	1484
	Qy	1605	TACGTTTAATGGTAGTGTACTTAAGTAGGGGTCTTAATTTTAAATACATCAATTCACAGG	1664
	Db	1485	AACAAGTGGAGGACCAATTCGCTTATACTATTGTTTAAATATAAATGGGCAATACCCCAAG	1544
	Qy	1665	GTATCGGCTGAGAGTTGTTATGCTGCTCTCAACAATGGTCTCGAGGGTAACTGTCGG	1724
	Db	1545	GTATCGTGGAAGAAATACGCTATCGCTCTACTACAAATCTAAGAAATTTAGTAAACGGTTGC	1604
	Qy	1725	AGGAGTACTACTTTTTGATCAAGGATTCCTCTAGTACTATGAGTGCAAAATGAGTCTTTGAC	1784
	Db	1605	AGGTGAACGGATTTTGTCTGTCNAATTTAAACAAAACANATGGATACGGGTGACCCATTAAAC	1664
	Qy	1785	ATCTCAATCATTTAGATTTCCGAAATTTCTCTGTTAGGTTATTAGTGCAATCTGGCAGTCAAAC	1844

Db 954 ACACCTAGTTAGTTCAACGAATACGGTGTAAACGGTATAAATTTCCCTAGTTACGGGGT 1013
Qy 1125 TACCATACTTCTATTAATCCTGTAAAGANTATCAATCTCTCTCGAGACGCTATATGGAC 1184
Db 1014 CTTCAATCCTGTGGCGCCATTTGGATTTGCAGATGAGGATCCACGTCCTTTTATCGGAC 1073
Qy 1185 TGAATCATATATCAGAGTGCTTCTATATGGGAATTTACCTTGAACCTATTATGTTGTCCTC 1244
Db 1074 ATTAATCAGATCCTGTTTTTGTCCGAGGAGGAT-----TTGGGAATCCTCATATGTAAT 1127
Qy 1245 TACTGTTAGATTTAAATTTTAGGAACCCCTCAGAAATACATTTTGAAGAGAGTACTGCTAACTA 1304
Db 1128 GGGGCTTAGGGAGTAGCATTTTCAACAACTGTGTACGNAACACACCCGAAATTTAGAAA 1187
Qy 1305 TAGTCAACCCCTATGAGTACCTGGCTTCAATTAAGAAATTCAGAAATCTGAATTTACCAAC 1364
Db 1188 TAGTGGGACCATGATTTCTCTAGATGAATCCCACTCAGGATAATAGTGGGGCACC- TT 1246
Qy 1365 AGAAACAAACAGAACGACCAAAATTAATGAATCATATAGTATAGTGTATCTCACATAGGGCT 1424
Db 1247 GGAATGATTAATGTCATGTATTAATCATGTTACATTTGTACGATGGCCAGGTGAGATTT 1306
Qy 1425 CATTTCAATCTAGGTGCAATGACCAATATTTCTTGGACGACCGTAGTGCAATCG 1484
Db 1307 CAGGAAGTGATTCATGAG- -AGCTCCAATGTTTTCTTGGACGACCGTAGTGCACCC 1364
Qy 1485 TACAAATACCAATTAGTTAGATAGCATAAACAAATACCAATGGTAAATCATTCAACCT 1544
Db 1365 TACAAATACCAATGATCCGAGAGGATTAATCAAAATACCAATGGTAAAGCACATACACT 1424
Qy 1545 TAATTCAGGTACCTCTCTAGTTCAGTGGCCAGGAAATTTACAGGAGGGGATATAATCCGMA 1604
Db 1425 TCAGTCAGGTACTACTGTTAGAGAGGCGCGGGTTTACGGAGGAGATATCTTCGACG 1484
Qy 1505 TAAGCTTAATGTAGTGTATTAAGTATGGTCTTAATTTTAAATTAATACATCATACAGCG 1664
Db 1485 AACAAAGTGGAGGACCAATTTGCTTATACTATTTGTTAATATAATTAATGGGCAATTTACCCCAAAG 1544
Qy 1665 GTATCGGTGAGAGTTCGTTATGCTGCTTCTCAACAAATGGTCTCGAGGGTAACTGTGG 1724
Db 1545 GTATCGTGAAGAAATACGCTATGCTCTACTACAAATCTAAGAAATTTACGTAAACGGTTGC 1604
Qy 1725 AGGAGTACTACTTTTGATCAAGGATTCCTCTAGTACTATGAGTGCAAATGAGTCTTTGAC 1784
Db 1605 AGGTGAACGGATTTTCTGCTGCTCAATTTAACAACAATGATACCGGTGACCCATTAAC 1664
Qy 1785 ATCTCAATCATTTAGATTTCAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAA 1844
Db 1665 ATTCCAATCTTTAGTTTACGCAACTATTAATACAGCTTTTACATTCCCAATGAGCCAGAG 1724
Qy 1845 TGCT- -GGAAATAGTATAAGTAAATATGCAAGGTAGACAAACGTTTCACTTTTGATAAAT 1901
Db 1725 TAGTTTACAGATAGGTGCTGATCTTTTAGTTTACGGGAATGAAGTTTATATAGACAGATT 1784
Qy 1902 TGAATTCATTTCCAAATTAATCTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGA 1961
Db 1785 TGAATTTGATTTCCAGTTACTGCAACATTTGAAGCAGAAATATGATTTTAGAAGAGCACAAA 1844
Qy 1962 GCGGTGGAATGCTCTGTTTACTTAATAGGAATCCAGAAATTTGAAACAGATGTGACAGA 2021
Db 1845 GCGGTGGAATGCGCTGTGTTTACTTCTATAAACCAAAATAGGGATTAACACAGATGTGACGGA 1904
Qy 2022 TTATCATATTTAGTCAAGTATCCAAATTTAGTGGCGTGTATTCGATGAATTTCTGCTTGA 2081
Db 1905 TTATCATATTTGATCAAGTATCCAAATTTAGTGGATTTGTTTATCAGATGAATTTTGTCTGGA 1964
Qy 2082 TGAAGAGAGAAATTAATGAGAAAGTGAATATGCGAAACGACTCAGTGTGATGAAGAAA 2141
Db 1965 TGAAGAGCGAGAAATTTGTCGAGAGAAATGCAACATGCGAAGCGACTCAGTGTGATGAGCGGAA 2024
Qy 2142 CTTTACTCCAGAGTCCAAATCTCATCTCAATTAAGCAACCACTTCAATATCTACTAA 2201
Db 2025 TTACTTCAAGATCCAAATCTCAAGGCATCAATAGGCAACTAGAC----- 2070

Qy 2202 TGAGCAATCGAATTTTCACTATCTATCCATGAACAACTCTGAACATGGATGGTGGGAAGTGA 2261
Db 2071 -----CGTGGTTGGAGAGAGTAC 2090
Qy 2262 GAACATTACATCCAGAGAGGAATGACGTATTTTAAAGAGAAATTAACGTCACTACCGGG 2321
Db 2091 GGATATTAACCATCAAAGAGGAGATGACGTATTTCAAAGAAAAATATGTCACTACCCAG 2150
Qy 2322 GACTTTTAATAGTGTATTCGACGTATTTATATCAAAAAATAGGAGAGTCCGAAATTTAA 2381
Db 2151 TACCTTTGATGAGTGTATTCGAACGTATTTATCAAAAAATAGATGAGTCCGAAATTTAA 2210
Qy 2382 AGCTTATACCTCGTACCAATTAAGAGGGTATATTTGAAGATAGTCAAGATTTTAGAGATATA 2441
Db 2211 ACCCTATACCTCGTTATCAATTAAGAGGGTATATCGAGGATAGTCAAGACTTTAGAAATCTA 2270
Qy 2442 TTTGATTCGTTATTAATGCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATG 2501
Db 2271 TTTGATCCCGCTATAATGCAAAAACAGAAAACAGTAAATGTGCTAGGTACGGGTTCTTTATG 2330
Qy 2502 GCGCTTTTCAAGTTCGAAAGCCCAATCGAAAGGTGCGAGAACCGAAATCGATGCGCACACA 2561
Db 2331 GCGCTTTTCAAGTTCGAAAGTCCCAATCAGAAAGTGTGGAGAACCGAATCGATGCGCGCAC 2390
Qy 2562 TTTTGAATGGAATCTCGATCTAGATTTGTTCTGCGAGAGATGGAGAAAAATGTGCGCATCA 2621
Db 2391 CCTTGAATGGAATCTCGATCTAGATTTGTTCTGCGAGAGACGGGAAAAATGTGCACATCA 2450
Qy 2622 TTCCCATCATTTCTCTTTGGATATTTGATATTCGATGCGACAGACTTGCATGAGATCTAGG 2681
Db 2451 TTGCGCATCATTTCTCTTGGACATTTGATGTTGGATGTACAGACTTTAAATGAGGACTTTGA 2510
Qy 2682 CGTGTGGTGTGTTTCAAGATTAAGACGCGAGAAAGTCTATGCAAGACTAGGGAATCTCGA 2741
Db 2511 TGTATGGTGTATTTCAAGATTAAGACGCAAGATGGCCATGCAAGACTAGGAAATCTAGA 2570
Qy 2742 ATTTATTTGAAGAGAAACCAATTTATAGGAGAACACTGTCTCGTGAAGAGAGCAGAGAA 2801
Db 2571 GTTTCTCGAAGAGAAACCAATTTAGTGGGGAAGCACTAGCTCGTGTGAAAGAGCAGAGAA 2630
Qy 2802 AAAATGGAGAGACAAACGTTGAAAAACTACAAATTTGAAACAAACAGGATTTATACAGAGC 2861
Db 2631 AAAATGGAGAGATAAACGTTGAAAAATTTGGAATTTGGAACAAATATTTGTTTAAAGAGC 2690
Qy 2862 AAAAGAGCTGTGATGCTTTATTTGTAGATTTCTCAATATATAATAGATTTCAAGCGGATAC 2921
Db 2691 AAAAGAACTGTAGATGCTTTATTTGTAACCTCTCAATATGATCAATTTACAAGCGGATAC 2750
Qy 2922 AAACATTTGCGATGATTCATCGGCGAGATAAACTTGTTCATCGAATTCGAGAGCTTATCT 2981
Db 2751 GAATATTTGCGATGATTCATGCGGCGAGATAAACTGTTTCATAGAAATTCGGAAGCGGTATCT 2810
Qy 2982 GTCAGAAATTTATCTGTTATCCCGGGTGTAAATTCGCGAAAATTTTGAAGAAATTAAGAAGTCTG 3041
Db 2811 TCCAGAGTTATCTGTGATTCGCGGTGTAAATGTAGACATTTTCGAAGAAATTTAAAGGGCG 2870
Qy 3042 CATTTATCATCTGCAATCTCCCTATACGATGCGAGAAATGTGCTTAAAAATTTGTTGATTTTAA 3101
Db 2871 TATTTTCACTGCAATCTTCTCTATATGATGCGAGAAATGTCAATTAATAAAACCGGTGATTTCAA 2930
Qy 3102 TAAATGATTTAGCATCTCGAATGTAAAGGGCATGTAGATGT- -ACAAACAGAGCCATCA 3158
Db 2931 TAAATGCTTATCATGCTGGAACGTGAAAGGGCATGTAGATGTAGAAAGCAAAAAACAACA 2990
Qy 3159 CCGTTCTGCTCTGTTTATCCAGAAATGGGAAGCAGAAAGTGTCAACAGCAGTTTCGGCTCTG 3218
Db 2991 CCGTTCCGCTCTGTTGTTCCGGAATGGAAGCAGAAAGTGTCAACAAAGTTCGTGTCTG 3050
Qy 3219 TCCGGGGCGTGGCTATATCTCCGTGTCACAGCGTACAAAGGGATATGAGAGGGGTTG 3278
Db 3051 TCCGGGCTGTGGCTATATCTCTTCTGTTGTCAAGCGGTACAAAGGGGATATGAGAGAGGTTG 3110

QY 3279 TGTAAAGATCCATGAAATCGAGAACAAATACAGACGAATCTAAATTTTAAATCTGTGAAGA 3338
DB 3111 CGTAACCAATTCATGAGATCGAGAACAAATACAGACGAATCTAAATTTTAAATCTGTGAAGA 3170
QY 3339 AGAGGAAGTGTATCCAAACGATACAGGAACGTTGTAATGATTTATCTGACACCAAGGTAC 3398
DB 3171 AGAGGAAGTGTATCCAAACGATACAGGAACGTTGTAATGATTTATCTGACCAATCAAGAAGA 3230
QY 3399 A-----GCAGTATGTAATTCCTCCGTAATGCTGGATATGAGGATGCATATGAATTTGATAC 3452
DB 3231 ATACGGGGTGGTACACTCTCCGTAATGCTGGATATGAGGATGCATATGAATTTGAGCAATTC 3290
QY 3453 TACAGATCTGTGTAATTAACAAACCGACTTATGAAGAGAAACGTTATACAGATGTACGAAG 3512
DB 3291 TTCTGTACCAGCTGATTAATGCTGAGTCTATGAGGAAATCGTATACAGATGGACGAAG 3350
QY 3513 AGATTAATCATCTGTAATATGACAGAGGGTATGTAATTAATCCACCACTACCACTGGTTA 3572
DB 3351 AGACAATCTCTGTAATCTTAAACAGAGGATATGGGATTAACACCACTACCACTGGCTA 3410
QY 3573 TATGACAAAGAAATAGAAATCTTCCAGAAACCGATAGGATGATGATTTGAGATTGGAGA 3632
DB 3411 TGTGACAAAGAAATAGAGTACTTCCAGAAACCGATAGGATGATGATTTGAGATTGGAGA 3470
QY 3633 AACGAAAGGAAATTTATTTGTAGACAGCGTGGAAATTTACTCTTTATGGAGGAA 3684
DB 3471 AACGAAAGGAAATTTATTTGTAGACAGCGTGGAAATTTACTCTTTATGGAGGAA 3522

RESULT 9

US-10-825-751-7
; Sequence 7, Application US/10825751
; Publication No. US20040194165A1
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: Novel Bacillus thuringiensis Isolate Active Against Lepidopteran
; TITLE OF INVENTION: Pests, and Genes Encoding Novel Lepidopteran-Active Toxins
; FILE REFERENCE: MA-43CDP2D4
; CURRENT APPLICATION NUMBER: US/10/825,751
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 09/837,961
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/521,344
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 08/933,891
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: US 08/356,034
; PRIOR FILING DATE: 1994-12-14
; PRIOR APPLICATION NUMBER: US 08/210,110
; PRIOR FILING DATE: 1994-03-17
; PRIOR APPLICATION NUMBER: US 07/865,168
; PRIOR FILING DATE: 1992-04-09
; PRIOR APPLICATION NUMBER: US 07/451,261
; PRIOR FILING DATE: 1989-12-14
; PRIOR APPLICATION NUMBER: US 371,955
; PRIOR FILING DATE: 1989-06-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 3522
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-825-751-7

Query Match 38.2%; Score 1408.4; DB 18; Length 3522;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1061; Indels 108; Gaps 9;

QY 285 TCAGTGGGAAATTTCTTAGACATGTCGAAACAACTTATAATCAACAATTAACGAAAA 344
DB 207 TGATTGGAGCTTATTTCTTTTACAGATTGAACAAATTTGATTGACCAAGAAATAGAACATT 266

QY 345 TGCTAGGAATACGGCACTTGTCTCGATTACAAAGTTTAGAGATTCTTTTAGAGCCTATCA 404
DB 267 GGAAGGAACCGGCAATTTACTACATTACGAGGGTTAGCAGATAGCTATGAATTTATAT 326
QY 405 ACAGTCACTTGAAGATTGGCTTAGAAAACCGTGTATGATGCAAGAAACGAGAAAGTGTCTTTA 464
DB 327 TGAAGCACTAAGAGAGTGGGAAGCAAAATCTTAATAATGACAAATTAAGGGAAGATGTGG 386
QY 465 TACCCAAATATATAGCTTTAGAACTTTGATTTTCTTAATCGGATGCCGCTTTTCGCAATTAG 524
DB 387 TATTGATTTGCTTAATACAGACGAGCTTTTAATAACAGCAATAAATAATTTTACACTTAC 446
QY 525 AAACCAAGAGTCTCCATTATTAATAGTATATGCTCAAGCTGCAAAATTTTACACTTATTT 584
DB 447 AAGTTTTGAATCCCTCTTTTATCGTCTATGTTCAAGGGCGAATTTTACATTATCACT 506
QY 585 ATTGAGAGATGCTCTCTTTTGTGTAGTAATTTGGGCTTATCATCGCAGGAAATTTCAAG 644
DB 507 ATTAAGAGACGCTGTATCGTTTGGGCAAGGTTGGGCACTGGATATAGCTACTGTGTTAATA 566
QY 645 TTATTATGAGCGCCAAAGTGGAAACAAACGAGAGATTATTCGCACTATTCGCTAGTAATGTA 704
DB 567 TCATTATTAATAGATTAATAAATCTTTATCATAGATATACGAAACATTTGTTGGACACATA 626
QY 705 TAATACAGCTCTAAATAGCTTTGAGAGGACAAATGCCCAAGTTGGGTGCGTTATAATCA 764
DB 627 CAATCAAGGATTAGAAAACTTAAGAGGTACTAATACTCGACAAATGGGCAAGATTCAATCA 686
QY 765 ATTCCGTAGAGATCTAACGTTTAGGGTATTTAGATCTAGTGGCACTATTTCCCAAGCTATGA 824
DB 687 GTTTAGGAGAGATTTAAACACTTACTGTATTAGATATCGTCTCTTTTTCGAACTACGA 746
QY 825 CACTCGCACTTATCCAAATAATACAGGTGCTCAGTTACACAGGGAAGTTTATACAGCGC 884
DB 747 TGTAGAACATATCCAAATTCAAACGCTCATCCAAATTAACAAGGGAATTTTATACAGTTTC 806
QY 885 AATTGGAGCAACAGGGGTAAATATATGGCAAGTATGAATTTGGTATAATAATAATATGCACTTC 944
DB 807 AGTAATTGAGGATTCTCCAGTTTCTGC-----TAAATATACCTAA 845
QY 945 GTTTTCCGCTATAGAGACTGCGGTTATCCGAAGCCCGCATCTACTGTGATTTTCTAGAAC 1004
DB 846 TGGTTTTAATAGGCGGAAATTTGGAGTTAGACCGCCCATCTTATGGACTTTATGAA--- 902
QY 1005 ACTTACAATTTTATGACACTTCATCCAGTGCAGTGTCTACTAGGCATATGACTTACTGCG 1064
DB 903 -----TTCTTTGTTGTAACCTGCAGAGACTGTTAGAAAGTCAAACTGTGTGGGAGG 953
QY 1065 GGGGCACACAATTCAAATCTCGGCCAATAGGAGGGGATTAATAATACCTCAACGCGATGGGTC 1124
DB 954 ACACCTTAGTTAGTTTCAAGAAATACGGCTGGTAAACCGTATAAATTTCCCTAGTTACGGGT 1013
QY 1125 TACCAATCTTCTATTAATCTCTGTAAGATTATCAATTTCTCTCGAGACGTTATTTGAC 1184
DB 1014 CTTCAATCCTGCTGGCGCAATTTGGATTTGAGATGAGGATCCAGCTCTCTTTTATCGGAC 1073
QY 1185 TGAATCATATCGCAGAGTGTCTTCTATGGGAATTTACCTGCACTATTTACTGTTGTC 1244
DB 1074 ATTATCAGATCTCTGTTTGTTCGAGGAGGAT-----TTGGGAATCTCTATTATGTACT 1127
QY 1245 TACTGTTAGATTTAATTTTAGGAACCTCTCAGAATACTTTTGAAGAGGATCTGCTAACTA 1304
DB 1128 GGGCTTAGGGGAGTAGCATTTTCAACAACTGGTACGAAACCCACCCGAAACATTTAGAAA 1187
QY 1305 TAGTCAACCTTATGAGTCACTGGGCTTCAATTTAAAGATTTCAGAACTGAATTTACCACC 1364
DB 1188 TAGTGGGACCATAGATTTCTCTAGATGAATAATCCACCTCAGGATAATAGTGGGCACT- 1246
QY 1365 AGAACACACGAAACGACCAATTTATGATCATATAGTATAGTATCTCATAGGGCT 1424
DB 1247 GGAATGATTTATGATCATGTTTAAATCATGTTTAAATGATGTTTACGATGGCCAGGTGAGATT 1306
QY 1425 CATTTCAATCTAGGGTGCATGTACCAAGTATATTTCTTGGACGCAACCGTAGTCAGATCG 1484

Db 1307 CAGGAAGTGATTCATGGAG--AGCTCCAATGTTTTCTTGGACGCCCGTAGTGCACACCCC 1364
Qy 1485 TACAATATACATTTAGTTTCAGATAGCATAAACACAAATACCATTGGTAAATATCAATCAACCT 1544
Db 1365 TACAATATACAAATTTGATCCGAGAGGATTAATCAATACCATTGGTAAAGACATACACT 1424
Qy 1545 TAAATTCAGGTACCTCTGTAGTCAGTCGCCCGAGGATTTACAGAGGGGATATTAATCCGAAC 1604
Db 1425 TCAGTCAGGTACTACTGTTGTAAAGAGGCCCGGGTTTACGGAGAGAGATATTTCTTCGACG 1484
Qy 1605 TAACGTTAATGGTGTACTAAGTATGGGTCTTAATTTTAATAATACATCAATTAACACGG 1664
Db 1485 AACAACTGGAGGACCAATTTGCTTATATCTATTTAATAATTAATTAATTAATTAATTAAT 1544
Qy 1665 GTATCCGCTGAGAGTTGCTTATGCTGCTCTCAACAATATGCTTCGAGGGTAACTGTGCG 1724
Db 1545 GTATCGTGCAGAAATACGCTATGCTCTACTACTCAAAATCTAAGAAATTTACGTTAAACGGTTGC 1604
Qy 1725 AGGAGTACTACTTTTGTATCAAGGATTCCTAGTACTATGATGCAAAAGCTTACATTTGATTAAT 1784
Db 1605 AGGTGAACGGATTTTGTGCTGGTCAATTTAAACAAACATGGATACCGGTGACCCATTAAC 1664
Qy 1785 ATCTCAATCATTTAGATTTGCGAATTTCTCTGTAGGTATTAAGTGCATCTGCGAGTCAAAC 1844
Db 1665 ATTCCAATCTTTTAGTACGCAACTATTAATACAGCTTTTACATTTCCCAATGAGCCAGAG 1724
Qy 1845 TGCT---GGAATAAGTATAAGTAATAATGTCAGGTAGACAAAGCTTACATTTGATTAAT 1901
Db 1725 TAGTTTCACAGTAGTGTGATCTTTTAGTTTCAGGGAATGAAGTTTATATAGACAGATT 1784
Qy 1902 TGAATTCATTCGAATTTACTGCAACTTCGAAAGCAGAAATACGATTTAGAAAGGGCGCAAGA 1961
Db 1785 TGAATTTGATTCGAGTTACTGCAACTTTGAAAGCAGAAATATGATTTAGAAAGAGACACAAA 1844
Qy 1962 GCGGTGAATGCTCTGTTTACTTAATACGAATCCAGAGATTTGAACACAGATGTGACAGA 2021
Db 1845 GCGGTGAATGGCTGTTTACTCTTAACCAATAGGGAATGAACACAGATGTGACGGA 1904
Qy 2022 TTATCATATTCATCAAGTATCCAAATTTAGTGGCGTGTATTCGGAATGAAATTCGCTTAGA 2081
Db 1905 TTATCATATTCATCAAGTATCCAAATTTAGTGGATGTTTATCAGATGAATTTTGTCTGGA 1964
Qy 2082 TGAAGAAGAGAAATTTACTTGAGAAAGTGAATATGCGAAACGATCAGTATGAAGAA 2141
Db 1965 TGAAGAAGAGAAATTTCTCGAAGAGTCAACATGCGAAGCGACTCAGTATGAGCGGAA 2024
Qy 2142 CTTACTCCAGATCCAAACTTCACATCCATCAATTAAGCAACACAGACTTCATATCTACTAA 2201
Db 2025 TTTACTTCAAGATCCAAACTTCAAAGGCATCAATAGGCAACTAGAC----- 2070
Qy 2202 TGAGCAATCGAATTTACATCTATCCATGAACAACTCTGAACATGGATGGTGGGAGTGGA 2261
Db 2071 -----CGTGGTTGGAGAGGAGTAC 2090
Qy 2262 GAACATTTAATCCAGGAAGAAATGACGTATTTAAAGAGAAATACGTACACACTACCGGG 2321
Db 2091 GGATATTAACCATCCAAAGAGAGATGACGTATTTCAAGAAATATTTGTACACACTACCAGG 2150
Qy 2322 GACTTTTAAAGTGTATCCGACGTATTTATATCAAAAAAATAGGAGAGTCCGGAATTTAA 2381
Db 2151 TACCTTTGATGAGTGTATCCAAACGTATTTATATCAAAAAAATAGATGAGTCCGAATTTAA 2210
Qy 2382 AGCTTATCTCGCTACCAATTAAGAGGGTATATGAAGATGCTCAAGATTTAGAGATATA 2441
Db 2211 ACCCTATCTCGTTATCAATTAAGAGGGTATATCGAGGATGCTCAAGACTTAGAAATCTA 2270
Qy 2442 TTTGATTCGTTATTAATCCGAAACATGAACATTTGATGTTTCCAGGTACCGATCCGATG 2501
Db 2271 TTTGATCCCGCTATTAATGCAAAAAACAGAAACAGTAATAATGTCTAGGTACCGGTTCTTTATG 2330
Qy 2502 GCCGCTTTTCAGTTGAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATCGGACCCACA 2561
Db 1307 CAGGAAGTGATTCATGGAG--AGCTCCAATGTTTTCTTGGACGCCCGTAGTGCACACCCC 1364
Qy 1485 TACAATATACATTTAGTTTCAGATAGCATAAACACAAATACCATTGGTAAATATCAATCAACCT 1544
Db 1365 TACAATATACAAATTTGATCCGAGAGGATTAATCAATACCATTGGTAAAGACATACACT 1424
Qy 1545 TAAATTCAGGTACCTCTGTAGTCAGTCGCCCGAGGATTTACAGAGGGGATATTAATCCGAAC 1604
Db 1425 TCAGTCAGGTACTACTGTTGTAAAGAGGCCCGGGTTTACGGAGAGAGATATTTCTTCGACG 1484
Qy 1605 TAACGTTAATGGTGTACTAAGTATGGGTCTTAATTTTAATAATACATCAATTAACACGG 1664
Db 1485 AACAACTGGAGGACCAATTTGCTTATATCTATTTAATAATTAATTAATTAATTAATTAAT 1544
Qy 1665 GTATCCGCTGAGAGTTGCTTATGCTGCTCTCAACAATATGCTTCGAGGGTAACTGTGCG 1724
Db 1545 GTATCGTGCAGAAATACGCTATGCTCTACTACTCAAAATCTAAGAAATTTACGTTAAACGGTTGC 1604
Qy 1725 AGGAGTACTACTTTTGTATCAAGGATTCCTAGTACTATGATGCAAAAGCTTACATTTGATTAAT 1784
Db 1605 AGGTGAACGGATTTTGTGCTGGTCAATTTAAACAAACATGGATACCGGTGACCCATTAAC 1664
Qy 1785 ATCTCAATCATTTAGATTTGCGAATTTCTCTGTAGGTATTAAGTGCATCTGCGAGTCAAAC 1844
Db 1665 ATTCCAATCTTTTAGTACGCAACTATTAATACAGCTTTTACATTTCCCAATGAGCCAGAG 1724
Qy 1845 TGCT---GGAATAAGTATAAGTAATAATGTCAGGTAGACAAAGCTTACATTTGATTAAT 1901
Db 1725 TAGTTTCACAGTAGTGTGATCTTTTAGTTTCAGGGAATGAAGTTTATATAGACAGATT 1784
Qy 1902 TGAATTCATTCGAATTTACTGCAACTTCGAAAGCAGAAATACGATTTAGAAAGGGCGCAAGA 1961
Db 1785 TGAATTTGATTCGAGTTACTGCAACTTTGAAAGCAGAAATATGATTTAGAAAGAGACACAAA 1844
Qy 1962 GCGGTGAATGCTCTGTTTACTTAATACGAATCCAGAGATTTGAACACAGATGTGACAGA 2021
Db 1845 GCGGTGAATGGCTGTTTACTCTTAACCAATAGGGAATGAACACAGATGTGACGGA 1904
Qy 2022 TTATCATATTCATCAAGTATCCAAATTTAGTGGCGTGTATTCGGAATGAAATTCGCTTAGA 2081
Db 1905 TTATCATATTCATCAAGTATCCAAATTTAGTGGATGTTTATCAGATGAATTTTGTCTGGA 1964
Qy 2082 TGAAGAAGAGAAATTTACTTGAGAAAGTGAATATGCGAAACGATCAGTATGAAGAA 2141
Db 1965 TGAAGAAGAGAAATTTCTCGAAGAGTCAACATGCGAAGCGACTCAGTATGAGCGGAA 2024
Qy 2142 CTTACTCCAGATCCAAACTTCACATCCATCAATTAAGCAACACAGACTTCATATCTACTAA 2201
Db 2025 TTTACTTCAAGATCCAAACTTCAAAGGCATCAATAGGCAACTAGAC----- 2070
Qy 2202 TGAGCAATCGAATTTACATCTATCCATGAACAACTCTGAACATGGATGGTGGGAGTGGA 2261
Db 2071 -----CGTGGTTGGAGAGGAGTAC 2090
Qy 2262 GAACATTTAATCCAGGAAGAAATGACGTATTTAAAGAGAAATACGTACACACTACCGGG 2321
Db 2091 GGATATTAACCATCCAAAGAGAGATGACGTATTTCAAGAAATATTTGTACACACTACCAGG 2150
Qy 2322 GACTTTTAAAGTGTATCCGACGTATTTATATCAAAAAAATAGGAGAGTCCGGAATTTAA 2381
Db 2151 TACCTTTGATGAGTGTATCCAAACGTATTTATATCAAAAAAATAGATGAGTCCGAATTTAA 2210
Qy 2382 AGCTTATCTCGCTACCAATTAAGAGGGTATATGAAGATGCTCAAGATTTAGAGATATA 2441
Db 2211 ACCCTATCTCGTTATCAATTAAGAGGGTATATCGAGGATGCTCAAGACTTAGAAATCTA 2270
Qy 2442 TTTGATTCGTTATTAATCCGAAACATGAACATTTGATGTTTCCAGGTACCGATCCGATG 2501
Db 2271 TTTGATCCCGCTATTAATGCAAAAAACAGAAACAGTAATAATGTCTAGGTACCGGTTCTTTATG 2330
Qy 2502 GCCGCTTTTCAGTTGAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATCGGACCCACA 2561

Db 2331 GCCGCTTTCACTCCAAAGTCCAATCAGAAAGTGTGGAGAACCGAATCGATGCGCGCCACA 2390
Qy 2562 TTTTGAATGAATCCTGATCTAGATTGTTCTCTCAGAGATGGAGAAAAATGTGCGCATCA 2621
Db 2391 CCTTGAATGAATCCTGATCTAGATTGTTCTCTCAGAGACGGGAAAAATGTGCACATCA 2450
Qy 2622 TTCCCATCAATTTCTTTTGGATATTTGATATTTGGATGCACAGACTTGCATGAGATCTAGG 2681
Db 2451 TTCCCATCAATTTCTTTTGGACATTTGATGTTGGATGTACAGACTTAAATGAGAGACTTAGA 2510
Qy 2682 CGTGTGGTGGTATTCAGAAATTAAGACGCGAGGAAGGTCTATGCAAGACTTAGGGAATCTGGA 2741
Db 2511 TGTATGGGTGATATTCAGAAATTAAGACGCGAAGATGCCATGCAAGACTTAGGAAATCTAGA 2570
Qy 2742 ATTTATTTGAAGAGAAAACCATTTATAGGAGAGGACACTGTCTGTGTGAAGAGAGAGAGAA 2801
Db 2571 GTTTCTCGAAGAGAAAACCATTTAGTCGGGGAAGCAGCTAGCTCGTGTGAAAAGAGCAGAGAA 2630
Qy 2802 AAAATGGAGAGACAAACGTAAGAACTCAAAATTTGGAACCAAAACGAGTATATACAGAGGC 2861
Db 2631 AAAATGGAGAGATAAACGTTGAAATTTGGAATTTGGAACCAAAATATTTGTTTAAAGAGGC 2690
Qy 2862 AAAAGAGCTGTGGATGCTTTATTTTGTAGATTCTCAATATATAGATTACAAGCGGATAC 2921
Db 2691 AAAAGAACTCTGTAGATGCTTTATTTTGTAACTCTCAATATGATCAATTAACAGCGGATAC 2750
Qy 2922 AAACATTTGCATGATTCATGCGGACAGATAAACTTGTTCATCGAATTCGAGAGGCTTATCT 2981
Db 2751 GAATATTTGCATGATTCATGCGGACAGATAAAGCTGTTTATAGAAATTCGGGAAGCGTATCT 2810
Qy 2982 GTCAGAAATTTCTGTTATTCGCGGTGTAAATTCGCGAAATTTTGAAGAAATTAGAAGTCG 3041
Db 2811 TCCAGAGTTATCTGTGATTCGCGGTGTAAATGTAGACATTTTTCGAAGAAATTAAGAGGCG 2870
Qy 3042 CATTATCACTGCAATCTCCCTATACGATGCGAGAAATGCTGCTTAAATATGCTGATTTTAA 3101
Db 2871 TATTTTCACTGCAATCTTCTTATATGATGCGAGAAATGCTCAITTAACACGCTGATTTCAA 2930
Qy 3102 TAATGGATTAGCATCTCGAATTTAAAGGGCATGTAGATGT---ACAACAGAGCCATCA 3158
Db 2931 TAATGGCTTATCATGCTGGAACGCTGAAAGGGCATGTAGATGTAGAGAAACAAACCAACA 2990
Qy 3159 CCGTTCTCTCTCTGTTATTCGCCAGAAATGGAGAGAGAAAGTGTCAACAGCAGTTCGCGTCTG 3218
Db 2991 CCGTTCTCTCTCTGTTATTCGCCAGAAATGGAGAGAGAAAGTGTCAACAGAAAGTTCGTTCTG 3050
Qy 3219 TCCGGGGCGTGGCTATATCTCTCGTGTACAGCGGTACAAAGAGGGATATGGAGAGGTTG 3278
Db 3051 TCCGGGTGCTGGCTATATCTCTTCTGTGTACAGCGGTACAGAGGGATATGGAGAGGTTG 3110
Qy 3279 TGTAAACGATCCATGAAATCGAGAACATACAGAGCACTAAATTTTAAACCTGTGAAGA 3338
Db 3111 CGTAAACATTCATGATTCGAGAACATACAGAGCACTGAAGTTTAGCAACTGCGGTAGA 3170
Qy 3339 AGAGAAAGTGTATCCAAACGGATACAGAAACGTTGATGATTATCTGCACACCAAGGTAC 3398
Db 3171 AGAGAAAGTGTATCCAAACCAACCGGTAAACGTTGATGATTATCTGCAATCAAGAGAA 3230
Qy 3399 A-----GCAGTATGTAATTTCCCGTAAATGCTGATGATGAGGATGCATATGAAGTGTATAC 3452
Db 3231 ATACGGGGGTGCGTACACTTTCCCGTAAATCGTGGATATGACGAAACCTTATGGAAAGCAATTC 3290
Qy 3453 TACAGCATCTGTTTAAATTTACAAACCGACTTATGAAGAGAAACCGTATACAGATGTACGAG 3512
Db 3291 TTTCTGACCAGCTGATTTATGCGGTGAGTCTATGAAGAAATTCGTATACAGATGGAGCAAG 3350
Qy 3513 AGATAATCATTTGTAATATGACAGAGGATGTGAAATTTATCCACCTACCACTACAGCTGGTTA 3572
Db 3351 AGACAATCTTTGTAATTTAAACAGAGGATATGGGGATTTACACACCACTACCACTGCTGCTA 3410
Qy 3573 TATGACAAAGAAATTAGAATTAATCTTCCAGAAACCGAATAAGGTATGATGATGAGATTGGAGA 3632
Db 3411 TGTGACAAAGAAATTAGAGTACTTCCAGAAACCGGATAGGTATGGATTGAGATCGGAGA 3470

QY 3633 AACGGAAGGAAATTTATTGTAGACAGCGTGGAAATTTACTCTTATCGAGGAA 3684
|||||
Db 3471 AACGGAAGGAAATTTACTCTGTGACAGCGTGGAAATTTACTCTTATGAGGAA 3522
|||||

RESULT 10

US-09-873-873-25
; Sequence 25, Application US/09873873
; Patent No. US20020064865A1
; GENERAL INFORMATION:
; APPLICANT: Malvar, Thomas
; APPLICANT: Gilmer, Amy Jelen
; FILE OF INVENTION: Polynucleotide Compositions Encoding Broad-Spectrum S-Endotoxins
; FILE REFERENCE: MECO:210--2
; CURRENT APPLICATION NUMBER: US/09/873,873
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 09/253,341
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: US 08/922,505
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: US 08/754,490
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; NAME/KEY: CDS
; LOCATION: (1)..(3531)
US-09-873-873-25

Query Match 38.0%; Score 1402; DB 9; Length 3534;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

QY 285 TCAGTGGGAAATTTCTTAGACATGTCGAACAATTTAAATCAACAATAACAGAA 344
Db 213 TCNATGGGAGCGATTTCTTTGACAAATTTGAAGATTTAAACCAAGAAATAGAAGATT 272
QY 345 TGCTAGGAATACGGCACTTGTCTGATTAACAAGTTTATAGAGATTCCTTTAGAGCCTATCA 404
Db 273 CGCTAGGAACCAAGCAATTTCTAGATTAAGAGCAATCTTATCAATTTACGC 332
QY 405 ACAGTCACCTTGAAGATTTGCTAGAAACCGTGATGATGATCAAGAACGAGAGTGTCTTTTA 464
Db 333 AGAATCTTTTAGAGAGTGGGAAAGCAGATCTACTAATCCAGCATTTAAGAGAGAGATGCG 392
QY 465 TACCAATATATAGCCTTAGAATTTGATTTCTTAATCGATGCGCTTTTCGCAATTAG 524
Db 393 TATTCATTTCAATGATCAAGACAGTGGCCCTTACACCGCTATCTCTTTTGCAGTTCA 452
QY 525 AAACCAAGAAGTTCCATTTAATGGTATATCTCAAGCTGCATAATTTACACCTATTATT 584
Db 453 AAATTTCAAGTTCTCTTTTATCAGTATATGTTCAAGCTGCATAATTTACATTTATCAGT 512
QY 585 ATTGAGAGATGCTCTCTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACG 644
Db 513 TTTGAGAGATGTTTCAGTGTTTTGGACAAAGGTGGGGAATTTGATGCGCGCACTATCAATAG 572
QY 645 TTATTTATGCGCCAGTGGGAACAAACAGAGATTTATTCGACTATTTCGTAGATGTTA 704
Db 573 TCGTTATTAATGATTTAATAGGCTTTATTGGCAACTTATACAGATCATGCTGACGCTGTTA 632
QY 705 TAATACAGCTCTAATATAGCTTTCAGAGGAGCAAAATCCGCAAGTTTGGGTGCGTTATAATCA 764
Db 633 CAATACGGGATTAGAGCGTGTATGGGACCGGATTTCTAGAGATTTGATTAAGATATATCA 692
QY 765 ATTCGCTAGAGATCTTAAGCTTAGGGGTATTAGATCTAGTGGCATTATTTCCCAAGCTATGA 824
|||||

Db 693 ATTTAGAGAGAAATTAACAATACTGTATTAGATATCGTTTCTCTATTTCGAACTATGA 752
QY 825 CACTCGCACTTATCCAAATAAATACAGAGTCTCAGTTTAAACAAGGGAAGTTTATACAGAGC 884
Db 753 TAGTAGAAGCATTCOAATTCGAACAGTTTCCCAATTTAAACAAGAGAAATTTATACAAACC 812
QY 885 A--ATTGAGCAACACAGGGGTAAATATGCAAGTATGAATTGGTATATAATAATAATGCACCT 942
Db 813 AGTATTAGAAAAATTTTATGATGTAGTTTTCGAGGCTCGGCTCAGGSCATAGAAAGAAGTAT 872
QY 943 TCGTTTTCGCTATAGAGACTGCGGTATTCGAAGCCCGCATCTACTTCTGATTTCTTAGAA 1002
Db 873 TAGGAGTCCACATTTGATGGATATATCTTAAACAGTATACCATCTATACGGATGCTCATAG 932
QY 1003 CAACCTTACAAATTTTAGCACTTCAACAGATGGAGTGTACTAGGCATATGACTTACTGG 1062
Db 933 GGGTTATTATTATGTTGGTCAGGGCATCA-----AATAATGGCTTCTCTGTAGGGTT 983
QY 1063 CGGGGGCAACAATTTCAATCTCGGCCAATAGGAGCGGATTTAAATPACCTCAACGCATGGG 1122
Db 984 TTCGGGGCCAGAAATTCACCTTTCCGCTATATGGAACACTATGGGA-----ATGCA 1032
QY 1123 TCTACCAATATCTTCTTAAATCTCTGTAAAGATTATCATTTCTCTCTCGAGAGCTATATTGG 1182
Db 1033 GCTCCACAACAACGATTTGTTGCTCAACTAGGTGAGGCGGTATAGAACATTTATCGTCC 1092
QY 1183 ACTGAATCATATGAGGAGTCTTCTATGGGGAATTTTACCTTGAACCTATTTCATGGTCTC 1242
Db 1093 ACTTTATAGAGAACCTTTTAAATATAGGGAATAATAATCAACAACTATCTGTTCTTGAC 1152
QY 1243 CCTACTGTAGATTTAAATTTTAGAACCTCTCAGAAATCTTTTGAAGAGGTACTCTCTAAC 1302
Db 1153 GGGAC----AGAATTTGCTTATGGAACCTCTCTCAAAATTTGCCATCCGCTGTATACAGAA 1208
QY 1303 TATAGTCAACCTATAGTCACTCGGCTTCAATTTAAAGATTCAGAACTG-AAATTACC 1361
Db 1209 AAGCGGAACGATAGTTCGCTGGATGAATACCGCCACAGAAATAACAACGTCGCCACTAG 1268
QY 1362 ACCAGAACACACAGAACCAAAATTTATGAATCATATAGTATAGTTTCTCACAATAGG 1421
Db 1269 GCAAGGATTTAGTCATCGATTTAAGCCATGTTTCAATGTTTCGT-----CAGGCTTTAG 1322
QY 1422 GCTCATTTCAAACTTAGGGTGCATGTACCAAGTATATTTCTTGGAGCACCCTAGTGCAGA 1481
Db 1323 TAATAGTAGTGAAGTATAAAGAGCTCCAAATGTTTCTTGGAGCACCCTAGTGCAC 1382
QY 1482 TCGTCAAAATACCAATTTAGTTTCAAGATAGATTAACAACAATACCATTTGGTAAATCAATCAA 1541
Db 1383 CCTACAAATAACAATTTGATCCGAGAGGATTTACTCAAAATACCATTTGGTAAAGACATAC 1442
QY 1542 CTTTAAATTCAGGTACCTCTGTAGTCAAGTGGCCAGGATTTTACAGAGGGGATATAATCG 1601
Db 1443 ACTTCAGTCAAGTACTCTGTGTGAAGAGGCCCGGGTTTACGGGAGAGATTTCTTCG 1502
QY 1602 AACTAAGCTTAATGGTATGTACTAAGTATGGGTCTTTAAATTTTAAATATATACATTTACA 1661
Db 1503 ACGAAACAGTGGAGCACTTTGCTTATCTATCTTATGTTTAAATATAATGGCAATTTACCCA 1562
QY 1662 GCGGTATCGGTGAGAGTTTCTGTATGCTGCTTCTCAACAATGGTCTCGAGGGGTAACTGT 1721
Db 1563 AAGGTATCGTCAAGAAATACGCTATGCTCTACTACAAATCTAAGAAATTTACGTAACGGT 1622
QY 1722 CGGAGGAGTACTACTTTTGTATCAAGGATTCCTCTAGTACTATAGTGCATAATAGTCTTT 1781
Db 1623 TGCAGGTGAACGGATTTTGTGCTGATTTTAAACAACAATGGATACCGGTGACCCAT 1682
QY 1782 GACATCTCAATCTATTAGATTTGAGAAATTTCTCTGTAGGTATTAGTGCATCTGGCAGTCA 1841
Db 1683 AACATTCCAATCTTTTAGTTACGCACTTATTAATACAGCTTTTACATTTCCCAATGAGCCA 1742
QY 1842 AACTGCT---GGAAATAGTATAGTAAATATGAGGAGTAGACAACAGTTTCTCTTTGATAA 1898
Db 1743 GAGTAGTTTTCACAGTAGTGTGATCTTTTGTAGTTTCAGGGAATGAAGTTTATATAGACAG 1802

1899 AATTGAATTCATTCCAAATTAATCTGCAACCTTCGAGCAGAAATACGATTTAGAAAGGCGCA 1958
1803 ATTTGAATTTGATTTCCAGTTACTGCAACATTTGAAGCAGAAATATGATTTAGAAAGCACA 1862
1959 AGAGCGGTGAATCTCTGTTTACTAATACGAATCCAGNAGATTTGAAAACGAGATGTGAC 2018
1863 AAAGCGGTGAATGCGGTGTTTACTTCTATAAACCAATAGGAGATAGGATAGGATGAGC 1922
2019 AGATTATCATATTGATCAAGATCCAAATTTAGTGGGTGTTTATCGGATGAATTTCTGCTT 2078
1923 GGATTAATCATATTGATCAAGTATCCAAATTTAGTGGATTTGTTATCAGATGAATTTGTCT 1982
2079 AGATGAAGAGAGAGAAATTTACTTCGAAAGTGAATATGCGAAACGACTCAGTGATGAAG 2138
1983 GGATGAAGAGAGAGAAATTTGTCGAGAAAGTCAAAATGCGAAGCACTCAGTGATGAGC 2042
2139 AACTTACTCCAAAGATCCAAACTTCCATCCATCAATAGCAACAGCACTTCCATATCTAC 2198
2043 GAATTTACTTCAAGATCCAAACTTCAAGGCATCAATAGGCAACTAGAC----- 2091
2199 TAATGAGCAATCGAATTTACATCTCTCATGCAATCTGAACATGATGGTGGGGAAG 2258
2092 -----CGTGGTTGGAGGGAAG 2108
2259 TGAGAACATTTACAATCCAGGAAGGAAATGACGATTTTAAAGAGAAATACGTTCACTACC 2318
2109 TAGCGATATTACCATCCAAAGAGGAGATGACGATTTCAAAGAAATTTATGTCACACTACC 2168
2319 GGGGACCTTTAATGAGTGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCCGGAAT 2378
2169 AGGTACCTTTGATGAGTGTATCCAAACATATTTGTATCAAAAAATCGATGAATCAAAAT 2228
2379 AAAAGCTTATCTCGCTACCAATTAAGAGGATATATGAAGATAGTCAAGATTTAGAGAT 2438
2229 AAAAGCTTTACCGCTTATCAATTAAGAGGATATATGAAGATAGTCAAGATTTAGAAAT 2288
2439 ATATTGATTCGTTAATGCGAAACATGAACATTTGGATTTCCAGGTACCGAGTCCGT 2498
2289 CTATTAAATTCGCTACAAATGCAAAACATGAACATGTAATGTGCGACGTACGGGTTCTT 2348
2499 ATGGCGGCTTTCAAGTTGAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGCACC 2558
2349 ATGGCGGCTTTCAAGTTGAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGCACC 2408
2559 ACATTTGAATGGAATCTCGATCTAGATTTCTCGCAGAGATGGAGAAATGTCGCA 2618
2409 ACACCTTGAATGGAATCTCGATCTAGATTTCTCGTGTAGGATGGAGAAATGTCGCCA 2468
2619 TCATTCCCATCATTTCTTTGGATATTTGATTTGGATGACAGACTTGCATGAGAAATCT 2678
2469 TCATTCCCATCATTTCTTTGGATATTTGATTTGGATGACAGACTTAAATGAGGACCT 2528
2679 AGGCGTGGGTGTATTCAGATTTAAGACGCGAGGAGTCAATGCAAGACTAGGGAATCT 2738
2529 AGGTGTATGGGTGATCTTTAAGATTAAGACGCAAGATGGCAGCAAGACTAGGGAATCT 2588
2739 GGAATTTATTGAAGAGAACCATTTATAGAGAACGACTGTCTCGTGTGAAGAGCAGA 2798
2589 AGAGTTTCTCGAAGAGAACCATTTAGTAGAAGCGCTAGCTCGTGTGAAGAGCAGA 2648
2799 GAAAAATGAGAGACAAACGTCGAAATTTGGAATGGGAAACAAATATCGTTTAAAGA 2858
2649 GAAAAATGAGAGACAAACGTCGAAATTTGGAATGGGAAACAAATATCGTTTAAAGA 2708
2859 GGCAGAGAGCTGTGGATCTTTATTTGATTTCTCAATATATAGATTAACAAGCGGA 2918
2709 GGCAGAGAGCTGTGGATCTTTATTTGTAACCTCTCAATATGATCAATTACAGCGGA 2768
2919 TACAAACATTTGGCATGATTCATGCGGCAGATAAATCTTTTCATCAATTCGAGAGCTTA 2978
2769 TACGAATATTGCCATGATTCATGCGGCAGATAAATCTTTTCATCAATTCGAGAGCTTA 2828

2979 TCTGTCAGAATTTATCTGTATTCGGGTGTAAATGCGGAAATTTTGAAGAAATAGAGG 3038
2829 TCTCCCTGAGCTGTCTGTGATTTCCGGGTGTCAATGCGGCTATTTTGAAGAAATAGAGG 2888
3039 TCGCATTTATCACTCAATCTCCCTATATACGATGCGAGAAATGTCTGTTAAAAATGGTATTT 3098
2889 GCGTATTTTCACTGCAATCTCCCTATATGATGCGAGAAATGTCTTAAAAATGGTATTT 2948
3099 TAATAATGGAATTAGCATGCTGGAATGTAAAGGGCATGTAGATGT--ACAACAGAGCCA 3155
2949 TAATAATGGCTTATCTCTGCTGGAACGTGAAAGGCGCATGTAGATGTAGAGAAACAA 3008
3156 TCACCGGTTCTGCTCTGTTTATCCAGAAATGGAAGCAGAGTGTCAAGAGTTCGCGT 3215
3009 CCAACGTTCCGTTCTGTTGTTCCGGAATGGGAAGCAGAGTGTCAAGAAAGTTCGTGT 3068
3216 CTGTCCGGGCGTGGCTATATCTCCGTCGTACAGCGTCAAAAGAGGATATGGAGAGG 3275
3069 CTGTCCGGGCGTGGCTATATCTCTGTCGTACAGCGTCAAGAGGATATGGAGAGG 3128
3276 TTGTGTAAACGATCCATGAAATCGAGAAATCAATACAGACGAATCTGAAAGTCTGCGT 3188
3129 TTGTGTAAACGATCCATGAAATCGAGAAATCAATACAGACGAATCTGAAAGTCTGCGT 3188
3336 AGAAGAGGAGTGTATCCAAACGATACAGAAACGAACTTAAATTTTAAAACTGTGA 3335
3189 AGAAGAGGAGTGTATCCAAACGATACAGAAACGAACTTAAATTTTAAAACTGTGA 3395
3396 TACAGCAGTATGTAATTTCCGTAATGCTGGATATGAGGATGCATATGAAGTTGATCTAC 3455
3249 AGAATACGAGGTCGCT-----ACACTTCTCGTAATCGAGATATAACGAAGTCTCTTC 3302
3456 AGCATCTGTTAATTAACAAACCGACTTATGAAGAGAAACGATATACAGATGTACGAGAGA 3515
3303 CGTACCAGCTGATATGCGTCAGTCTATGAAGAAATTCGTATACAGATGGACGAGAGA 3362
3516 TAATCAATTTGTAATGACAGAGGATGTGAATTTATCCACACTTACCAGCTGTTATAT 3575
3363 GAATCTTTGTGAATTTAAGAGGATATAGGATTTACAGGCCACTACCAGTTGTTATGT 3422
3576 GACAAAGAAATTTAGAAATCTTCCAGAAACCGATAAGGTATGAGATTGAGATTGAGAAAC 3635
3423 GACAAAGAAATTTAGAAATCTTCCAGAAACCGATAAGGTATGAGATTGAGATTGAGAAAC 3482
3636 GGAAGGAAATTTATGATGACAGCGTGGAAATTTACTCTTTATGAGGAATAG 3687
3483 GGAAGGAAATTTATGATGACAGCGTGGAAATTTACTCTTTATGAGGAATAG 3534

RESULT 11

US-09-916-956A-25
; Sequence 25, Application US/09916956A
; Publication No. US20030017571A1
; GENERAL INFORMATION:
; APPLICANT: Malvar, Thomas
; APPLICANT: Gilmer, Amy Jelen
; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad-Spectrum S-Endotoxins
; FILE OF INVENTION: MECO-2111--1
; CURRENT APPLICATION NUMBER: US/09/916,956A
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 09/253,331
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: US 08/922,505
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: US 08/754,490
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

OTHER INFORMATION: Hybrid Delta-Endotoxin
; NAME/KEY: CDS
; LOCATION: (1) ... (3531)
US-09-916-956A-25

Query Match 38.0%; Score 1402; DB 10; Length 3534;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

QY	285	TCAGTGGGAAATTTCTTAGAACATGTCGAACAACTTATAATCAACAAATAACAGAAA	344
DB	213	TCATGGGACGATTTCTTGTACAAATGAACAGTTAATACCAAGAAATAGAAGATT	272
QY	345	TGCTAGGAATACGGCACTTGCTCGATTAACAAGTTTAGGAGATTCCCTTTAGAGCCATCA	404
DB	273	CGCTAGGAACCAAGCCATTCTTAGATTAGAAGCACTTAAGCAATCTTTATCAATTTACGC	332
QY	405	ACAGTCACTTGAAGATTGCTGAGAAAACCGTGATGATGCMAGACGAGAAGTGTCTTTA	464
DB	333	AGAATCTTTTAGAGAGTGGGAAGCAGATCTCTAATCCAGCATTAAGAGAGAGATGCG	392
QY	465	TACCAATATATAGCCTTAGAACCTTGATTTCTTAATGCGATGCCGCTTTTCGCAATTAG	524
DB	393	TATTCATTTCAATGATGACAGTGCCTTACACCGCTATTCCTCTTTTGGCAGTTCA	452
QY	525	AAACCAAGAAGTTCCATTTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATT	584
DB	453	AAATTTATCAAGTTCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTATCAGT	512
QY	585	ATTGAGAGATGCTCTCTTTTGGTGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACG	644
DB	513	TTTGAGAGATGTTTCAGTGTTTTGGACAAAGGTGGGGAATTTGATGCCGCACTATCAATAG	572
QY	645	TTATTATAGCGCCAGTGGACAAACAGAGATTATTCGCACTATTGGGTAGATGCTA	704
DB	573	TGTTTATATGATTTAATAGGCTTTATGGCAACTATACAGATCATGCTGTACGCTGGTA	632
QY	705	TAATACAGCTCTAAATAGCTTGAGAGGGACAAATGCCCAAGTTTGGGTGCGTTATPAATCA	764
DB	633	CAATACGGGATTAGACGGTGTATGGGACCGGATTTCTAGAGATTGGTAAGATATATCA	692
QY	765	ATTCGTTAGAGATCTAAAGTTTAGGGGTATATGATCTAGTGGCACTATTTCCCAAGCTATGA	824
DB	693	ATTTAGAAGAGAATTAACACTGATGTATAGATATCGTTTCTCTATTTCCGAACATGA	752
QY	825	CACTCGCACTTATCCAAATAATACAGTGCCTCAGTTTAAACAGGGAAGTTTATACAGACG	884
DB	753	TAGTAGAACGTATCCAAATTCGAACAGTTTCCCAATTAACAGAGAAATTTATACAAACCC	812
QY	885	A--ATTGAGACCAACAGGGTAAATATGCAAGTATGAATTTGGTATAATAATAATGCACCT	942
DB	813	AGTATTAGAAATTTTGTATGTTGATGTTTTCGAGGCTCGGCTCAGGCGCATAGAAGAGTAT	872
QY	943	TGTTTTCCGCTATAGAGACTGCGGTATACGAAGCCGCACTCTACTTCTGATTTCTAGAA	1002
DB	873	TAGGAGTCCACATTTGATGGATATATCTTAAACAGTATAACCATCTATACGGATGCTCATAG	932
QY	1003	CAACTTACAAATTTTAGCACTTCATCAGATGGAGTGTCTAGGCAATGACCTTACTTGG	1062
DB	933	GGTTTATTATTGGTCAGGCACTCA-----AATAATGGCTTCTCCTGTAGGGTT	983
QY	1063	CGGGGGCACAAATTTCAATCTCGGCAATAGGAGCGGATTAATAATCAACGATGGG	1122
DB	984	TTGCGGGCCAGNATTCATTTTCCGCTATATGGAACCTATGGAA-----ATGCA	1032
QY	1123	TCTACCAATATCTTATTAATCTGTAAAGATTATCAATCTTCTCTCGAGACGTATATGG	1182
DB	1033	GCTCCACAACAACGTAATTTGCTCAACTAGTACGGGCGTGTATAGAACATTTATCGTCC	1092
QY	1183	ACTGAATCATATGAGGAGTGTCTTATGCGGAATTTACCTTGAACCTTATTCATGGTCTC	1242
DB	1093	ACTTTATATAGAAGACCTTTTAAATATAGGGAATAATAATCAACACTATCTGTTCTTGAC	1152
QY	1243	CCTACTGTAGATTTAATTTTAGAACCCCTCAGAAATACTTTTGAAGAGGACTACTGTAAC	1302
DB	1153	GGGAC----AGATTTGGCTTATGGACCTCTCAAAATTTGCCATCCGCTGTATACAGAAA	1208
QY	1303	TATAGTCAACCCCTATAGTCACTCGGGCTTCAATTTAAAGAAATTCAGAACTG-AATTAAC	1361
DB	1209	AAGCGGAACGGTAGATTGCTGGATGAATACCGCCACAGAAATCAACACGTCGCCACTAG	1268
QY	1362	ACCAGAAACAACAGAACGACCAAAATTAATGAATCATATATAGTCTATAGTTATCTCACATAGG	1421
DB	1269	GCAAGATTTATAGTCATCGAATTAAGCCATGTTTCAATGTTTCGT-----TCAGGCTTTAG	1322
QY	1422	GCTCATTTCACAATCTTAGGGTGCATGTACCACTATATTCTTTGGACGACCGTAGTGCAGA	1481
DB	1323	TAATAGTAGTGTAAAGTATAAAGAGCTCCAAATGTTTCTTTGGACGACCGTAGTGCAC	1382
QY	1482	TCGTACAAATACCAATTTAGTTTCAGATAGCATTAACAAATACCAATTTGGTAAATTCATCOA	1541
DB	1383	CCCTACAAATCAAAATTTGATCCGGAGAGGATTTACTCAATACCAATTTGGTAAAGCACATAC	1442
QY	1542	CCTTAATTCAGGTACCTCTGTAGTCAGTGGCCCAAGGATTTTACAGAGGGGATATAATCCG	1601
DB	1443	ACTTCAGTCAGTACTCTGTGTAAGAGGCCCGGGTTTACGGGAGAGATATTCCTTCG	1502
QY	1602	AACTAACGTTAATGGTAGTGTACTAAGTATGGGTCTTTAAATTTTAATATACATCATTTACA	1661
DB	1503	ACGAACAAGTCGGAGGACCAATTTGCTTATCTATCTATTGTTTAATATAATGGCAATTACCCCA	1562
QY	1662	GCGGTATCGCTGAGAGTTCGTTATGCTGCTTCTTCAAAACAATGGTCTCGAGGGTAACTGT	1721
DB	1563	AAGGTATCGTCAAGAAATACGCTATGCTCTACTACAAATCTAAAGAAATTTACGTAACGGT	1622
QY	1722	CGGAGGAGTACTACTTTTGATCAAGGATCCCTAGTACTATGATGTCGCAATGAGTCTTT	1781
DB	1623	TGCAGGTGAACGGATTTTGTGCTGCTCAATTTAAACAAACAATGGATACCGGTGACCCCA	1682
QY	1782	GACATCTCAATCATTTTAGATTTGAGAAATTTCTGTAGGTATTAGTGCATCTGGCAGTCA	1841
DB	1683	AACATTCCAATCTTTTAGTTAGCGCACTATTAATACAGCTTTTACATTTCCCAATGAGCCA	1742
QY	1842	AACCTGCT---GGAAATAGTATAAGTATAATATGCAAGGTAGACAAAGTTTCTCTTTGATAA	1898
DB	1743	GAGTAGTTTCAAGTAGTGTGATCTATTTAGTTTCAGGGAATGAAGTTTATATAGACAG	1802
QY	1899	AATTTGAATTCATTCCAATTTACTGCAACTTCGAGACGAGAAATACGATTTTAGAAAGGCGCA	1958
DB	1803	ATTTGAATTTGATCCAGTTACTGCAACATTTGAAGACGAGAAATGATGTTTAGAAAGACACA	1862
QY	1959	AGAGCGGTGAATGCTCTGTTTACTTAATAACGAATCCAAAGAGATTGAAAAACAGATGTGAC	2018
DB	1863	AAAGCGGTGAATCGGCTGTTTACTTCTATAAACCAATAGGGATAAAAACAGATGTGAC	1922
QY	2019	AGATTATCATATTGATCAAGTATCCAAATTTAGTGGCGTGTGTTTATCGGATGAATTTCTGCT	2078
DB	1923	GGATTATCATATTGATCAAGTATCCAAATTTAGTGGATGTTTATCAGATGAATTTTGTCT	1982
QY	2079	AGATGAAGAGAGAAATTTACTTTGAGAAAGTGAATATGCGAAACGACTTCAGTGTAGAAAG	2138
DB	1983	GGATGAAGAGAGAAATTTGTCGGAAGAGTCAACATGCGAAGCGACTCAGTGTAGAGCG	2042
QY	2139	AACTTACTCTCAAGATCCAACTTCACATCCATCAATAAGCAACACAGACTTCATATCTAC	2198
DB	2043	GAATTTACTTCAGATCCAACTTCAAAGGCACTCAATAGGCACTAGAC-----CGTGGTTGAGAGGAAG	2091
QY	2199	TAATGAGCAATTCGAATTTCACTCTATCCATGAACAATCTGAACAATGGATGGTGGGGAAG	2258
DB	2092	-----CGTGGTTGAGAGGAAG	2108
QY	2259	TGAGAACTTACATCCAGGAGGAATGACGTATTTAAAGAGAAATTTACGTCACACTACC	2318
DB	2109	TACGGATTTACCATCCAAAGAGAGATGACGTATTCAAAGAAATTTATGTGCACACTACC	2168
QY	2319	GGGACTTTTAAATGAGTGTATCCGAGCTATTTATATCAAAAAATAGGAGAGTCCGAATT	2378


```
Db 2169 AGGTACCTTTGATGAGTGCTATCAACATATTTGTATCAAAAAATCGATGAATCAAAAT 2228
QY 2379 AAAAGCTTATATCTCGCTACCAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGAT 2438
Db 2229 AAAAGCCTTTACCGGTTATCAATTAAGAGGGTATATCGAAGATAGTCAAGACTTTAGAAAT 2288
QY 2439 ATATTTGATTCGTTATATATCGGAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGT 2498
Db 2289 CTATTTAAATTCGCTACATATCAAAACATGAAACAGTAAATGTGCCAGGTACGGGTTCTTT 2348
QY 2499 ATGGCCGCTTTCAGTTGAAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGGCACC 2558
Db 2349 ATGGCCGCTTTCAGCCCAAGTCCAATCGGAAGGTGCGGAGAACCGAATCGATCGGCACC 2408
QY 2559 ACATTTTGAATGAAATCCCTGATCTAGATTTGCTTCGAGAGATGAGAAAAATGTGCGCA 2618
Db 2409 ACACCTTGAATGAAATCCCTGATCTAGATTTGCTTCGAGAGATGAGAAAAATGTGCGCA 2468
QY 2619 TCATTTCCCATCAFTTCTCTTTGGATATTCATATTTGGATGCAAGACTTGCATGAGATCT 2678
Db 2469 TCATTTCCCATCAFTTCTCTTTGGATATTCATATTTGGATGCAAGACTTGCATGAGATCT 2528
QY 2679 AGCGGTGTCGGTGTATTCAGATTAAGACGCGAAGGTCTATGCAAGCTAGGGAATCT 2738
Db 2529 AGGTGTATGGGTGATCTTTAAGATTAAGACGCGAAGGTCTATGCAAGCTAGGGAATCT 2588
QY 2739 GGAATTTATTTGAAGAGAACCAATTAATAGGAGNAGCACTGTCTGTTGAAGAGAGAGA 2798
Db 2589 AGAGTTTCTCGAAGAGAACCAATTAATAGGAGNAGCACTGTCTGTTGAAGAGAGAGA 2648
QY 2799 GAAAAATGAGAGAGACAAACGTTGAAAACTACAATTTGAAAAACAAACAGGTATATACAGA 2858
Db 2649 GAAAAATGAGAGAGACAAACGTTGAAAACTACAATTTGAAAAACAAATATCTGTTTAAAGA 2708
QY 2859 GGCAAAAGAGCTGTGGATGCTTTATTTGTAGATTTCAATATATATAGATTACAAGCGGA 2918
Db 2709 GGCAAAAGAGCTGTGGATGCTTTATTTGTAACTCTCAATATATATAGATTACAAGCGGA 2768
QY 2919 TACAACATTTGGCATGATTTATCGGCGAGATATAACTTGTTCATCGAATTCGAGAGGCTTA 2978
Db 2769 TACGAATATTGCCATGATTCATCGGCGAGATATAACTTGTTCATCGAATTCGAGAGGCTTA 2828
QY 2979 TCTGTCAGAATATCTGTATCCCGGTGTAAATGCGGAAATTTTGAAGAATTAGAAGG 3038
Db 2829 TCTGCTGAGCTGTGTGATTCGCGGTGTCAATGCGGCTATTTTGAAGAATTAGAAGG 2888
QY 3039 TCGCATATCACTGCAATCTCCCTATACGATGCGAGAAATGTCGTTAAATGCTGATTT 3098
Db 2889 CGGTATTTTCACTGCAATCTCCCTATATATGTCGAGNAATGTCAATTAATAATGCTGATTT 2948
QY 3099 TAATAATGGATTAGCATGCTGGAAATGTAAAGGGCATGTAGATGT--ACAACAGAGCCA 3155
Db 2949 TAATAATGGATTATCTCGCTGGAAACGTTGAAGGGCATGTAGATGTGAAGAACAACAA 3008
QY 3156 TCACCGTCTCTGCTTCTGTTATCCAGAAATGGAAGCAGAGGTGTCAACGAGCTTCCGT 3215
Db 3009 CCAACGTTCCGTCTCTGTTTCCGNAATGGGAAGCAGAGGTGTCAACAGNAATGCTGCT 3068
QY 3216 CTGTCCGGGCGGTGGCTATATCTCCGTGTACAGCGTTACAAAGGGGATATGAGAGAGG 3275
Db 3069 CTGTCCGGGCGGTGGCTATATCTCCGTGTACAGCGTTACAAAGGGGATATGAGAGAGG 3128
QY 3276 TTGTGTAAACGATCCATGAAATCGAGAACATACAGACGAATCAAAATTTAAACCTGTGA 3335
Db 3129 TTGGTAAACATTCATGAGATCGAGAACATACAGACGAATCAAAATTTAGCAACTGGCT 3188
QY 3336 AGAAGAGAGATGTATCCAAACGATACAGAAACGTCGTGTAATGATTAATCTGCAACCAAGG 3395
Db 3189 AGAAGAGAGATGTATCCAAACGATACAGAAACGTCGTGTAATGATTAATCTGTAATCAAGA 3248
QY 3396 TACAGCAGTATGTAAATCCCGTAAATGCTGGATATGAGGATGATATGAATGTGATCTACTAC 3455
```

```
Db 3249 AGAATACGAGGTCGGT-----ACATTTCTCGTAATCGAGGATATATACGAAGCTCCTTC 3302
QY 3456 AGCATCTGTTAAATTAACAAACCGACTTATGAAGAAGAAACGTATACAGATGTACGAAGAGA 3515
Db 3303 CGTACCAGCTGATTTATCGGTCAGTCTATGAAGAAATCGTATACAGATCGAAGAGA 3362
QY 3516 TAATCATTTGTAATATACAGAGGGTATGTGAATTTATCCACCACTACCACTGGTTATAT 3575
Db 3363 GAATCTTTGTAATTTAACAGAGGGTATAGGGATTAACGCCACTACCACTGGTTATGT 3422
QY 3576 GACAAAAGAAATTAGAATTAATTTCCAGAAACCGATTAAGGTATGATTTGAGATTGAGAAAC 3635
Db 3423 GACAAAAGAAATTAGAATTAATTTCCAGAAACCGATTAAGGTATGATTTGAGATTGAGAAAC 3482
QY 3636 GGAAGGGAAGTTTATTTGTAGACAGCGTGGAAATTAATCTCTTATGGAGGAATAG 3687
Db 3483 GGAAGGGAACATTTATCTGTGACAGCGTGGAAATTAATCTCTTATGGAGGAATAG 3534

RESULT 12
US-09-997-914-25
; Sequence 25, Application US/09997914
; Publication No. US20030119158A1
; GENERAL INFORMATION:
; APPLICANT: Malvar, Thomas
; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad Spectrum d-Endotoxins
; FILE REFERENCE: 11792.0215.DVUS01 MECO:215--1
; CURRENT APPLICATION NUMBER: US/09/997,914
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/261,040
; PRIOR FILING DATE: 1999-03-02
; PRIOR APPLICATION NUMBER: US 08/754,490
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; NAME/KEY: CDS
; LOCATION: (1)...(3531)
US-09-997-914-25

Query Match 38.0%; Score 1402; DB 10; Length 3534;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

QY 285 TCAGTGGGAATTTTCTCTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAAA 344
Db 213 TCAATGGGACGCAATTTCTGTACAAATTTGAACAGTTAAATTAACCAAGAATAGAAGATT 272
QY 345 TGCTAGGAATACCGCACTTCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTATCA 404
Db 273 CGCTAGGAACCAAGCCATTTCTAGATTAGAGCACTAAGCAATCTTTATCAAAATTTACGC 332
QY 405 ACAGTCACTTGAGATTGGCTAGAAAACCGTGATGATGCAAGAACGAGAGATGTTCTTTA 464
Db 333 AGAATCTTTTAGAGAGTGGGAGCAGATCCTACTAATCCAGCATTAAGAGAGAGATGCG 392
QY 465 TACCAATATATAGCCTTAGAACTTCGATTTCTTAATCGATGCCGCTTTTGCATTTAG 524
Db 393 TATTCATTCATGACATGACAGTCGCTTACACCGCTATTTCTCTCTTTTTCGAGTTCA 452
QY 525 AAACCAAGAAGTTCCATTTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTAT 584
Db 453 AAATATATCAAGTTCTCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTATCATCT 512
QY 585 ATTGAGAGATGCTCTCTTTTGGTAGTGAATTTGGGCTTACATCGCAGAAATTCACAG 644
Db 513 TTTGAGAGATGTTTCACTGTTTGGACAAAGGTTGGGATTTGATGCGCGACTATCAATAG 572
```



```
Db 2649 GAAAAATGGAGACAAACCTGAAAAATTTGGAATGGGAACAAATATCGTTTATAAGA 2708
QY 2859 GGCRAAAGAACTGTGGAGCTTTATTTGTAGATTTCTCAATATAATAGATTACAAGCGGA 2918
Db 2709 GGCRAAAGAACTGTAGATGCTTTATTTGTAACTCTCAATATGATCAATTAACAGCGGA 2768
QY 2919 TACAAACATTTGGCATGATTCATCGGCGAGATAAATCTGTTTCATCGAATTCGAGAGCTTGA 2978
Db 2769 TACGAATATTTGCCATGATTTATCGGCGAGATAAATCGTTTCATAGCAATTCGAGAAGCTTA 2828
QY 2979 TCTGTCAGAAATATCTGTTATCCCGGGGTAAATCGGAAATTTTGAAGAATTAGAAG 3038
Db 2829 TCTGCTGAGCTGTCTGTGATTCGCGGTGTCAATCGGCTATTTTGAAGAATTAGAAG 2888
QY 3039 TCGCAATATCACTGCAATCTCCCTATACGATCGGAAATGTCGTTTAAAAATGTTGATTT 3098
Db 2889 CGGTATTTTCACTGCAATCTCCCTATATGATGCGAATAATGTCATTAATAATGGTGAAT 2948
QY 3099 TAATAATGGATTAGCATGCTGGAATGTAAGGCGCATGTAGATGTC---ACAACAGAGCCA 3155
Db 2949 TAATAATGGCTTATCTGCTGGAACGTGAAGGCGATGTAGATGTAGAGAACAAACAA 3008
QY 3156 TCACCGTCTGCTGCTGTTTATCCGGAATGGGAAGCGAGAGTGTCAAGCAGTTCGCGT 3215
Db 3009 CCAACGTTTCGGTCTGTTGTTTCCGGAATGGGAAGCGAGAGTGTCAAGAAAGTTCGTT 3068
QY 3216 CTGTCGGGGCGTGGCTATATCTCGTGTACACGCTACAAAGGCGATATGAGAGGG 3275
Db 3069 CTGTCGGGGCTGGCTATATCTCGTGTACACGCTACAAAGGCGATATGAGAGGG 3128
QY 3276 TTGTGTAAACGATCCATGAATCGAACAATAACAGCAACTAAAACTGTGA 3335
Db 3129 TTGCGTAACCATTCATGATGATCGAACAATAACAGCAACTGAAGTTTACCACTGCGT 3188
QY 3336 AGAAGAGGAAGTGTATCCAAACCGACTATGAGAGAACGTTGAATGATTAATGCAACCAAG 3395
Db 3189 AGAAGAGGAATCTATCCAAATACACGCTAACGTTGAATGATTAATGCAACCAAG 3248
QY 3396 TACACGATGTAATTCCTGTAATGCTGGATATGAGGATGCAATATGAAGTGTGATCTAC 3455
Db 3249 AGAATACCGAGGTGCGT-----ACACTTCGTAATCGAGGATATACGAAGCTCTTTC 3302
QY 3456 AGCATCTGTTAATACAAACCGACTATGAGAGAACCGTATACAGATGTACGAAGAGA 3515
Db 3303 CGTACCAGCTGATATGCGTCACTATGAGAGAAAAATCGTATACAGATGGACGAAGAGA 3362
QY 3516 TAATCATTTGTAATATGACAGAGGTATGTGAATTTATCCACCACTACCACTGCTGTTAT 3575
Db 3363 GAATCTTGTGAATTAACAGAGGTATAGGGATTACACGCCACTACCACTGCTGTTATGT 3422
QY 3576 GACAAAGAAATTAGAACTATCCAGAACCGATAGGATGATGATGAGATTGGAGAAAC 3635
Db 3423 GACAAAGAAATTAGAACTATCCAGAACCGATAGGATGATGATGAGATTGGAGAAAC 3482
QY 3636 GGAAGGAAGTTTATTTAGACAGCGTGAATTTACTCTTATGAGGAATAG 3687
Db 3483 GGAAGGAACATTTATCTGAGACAGCGTGAATTTACTCTTATGAGGAATAG 3534
```

RESULT 13

```
US-10-365-645-25
; Sequence 25, Application US/10365645
; Publication No. US20030182682A1
; GENERAL INFORMATION:
; APPLICANT: Malvar, Thomas
; APPLICANT: Gilmer, Amy Jelen
; TITLE OF INVENTION: Antibodies Immunologically Reactive with Broad-Spectrum
; TITLE OF INVENTION: Delta-Endotoxins (Amended)
; FILE REFERENCE: 11792.0210.DVUS02 (MECO:210--3)
; CURRENT APPLICATION NUMBER: US/10/365,645
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: US 09/873,873
```

```
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 09/253,341
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: US 08/922,505
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: US 08/754,490
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; NAME/KEY: CDS
; LOCATION: (1)..(3531)
; US-10-365-645-25
```

```
Query Match 38.0%; Score 1402; DB 15; Length 3534;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

QY 285 TCAGTGGGAAATTTTCTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAA 344
Db 213 TCAATGGGACGCAATTTCTGTACAAATTTGAACAGTTAATTAACCAAGAAATGAAGAAT 272
QY 345 TGTCTAGGATACGGCACTGCTCGATTAACAAGTTTAGAGATTCCTTTAGAGCCTATCA 404
Db 273 CGCTAGAACCAAGCCATTTCTAGATTAGAAGCACTAAGCAATCTTTATCAATTTACGC 332
QY 405 ACAGTCACCTTGGAATTTGGTAGAAACCGTAGAAACCGTAGATGCAAGACGAGAAAGTGT 464
Db 333 AGAATCTTTTGAAGAGTGGAGCAGATCTCTAATCCAGCATTAAGAGAGAGATGCG 392
QY 465 TACCCAAATATAGCCTTAGAACTTGATTTCTTAATCGCATGCGCTTTTCGCAATTAG 524
Db 393 TATTCATTTCAATGACATGAACAGTGGCCCTTACAAACGCTATTCTCTTTTGCAGTTCA 452
QY 525 AAACCAAGAAAGTTCATTTAATGTTATGTTATGCTCAAGCTGCAATTTACACCTATTATT 584
Db 453 AAATTTATCAAGTCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTATCAGT 512
QY 585 ATTGAGAGATGCTCTCTTTTGTAGTGAATTTGGGCTTACATCGCAGGAAATTTCAACG 644
Db 513 TTTGAGAGATGTTTCACTGTTTGGACAAAGTGGGGATTTTGATGCCGCGACTATCAATAG 572
QY 645 TTATTATGAGCGCCCAAGTGGAAACAAACGAGAGATTTATCCGACTATTTCGCGAATGGTA 704
Db 573 TCGTTAATAATGATTTAATAGGCTTATGGCAACTATACAGATCATGCTGTACGCTGGTA 632
QY 705 TAATAAGGCTCTAAATAGCTTGGAGGGAACAAATGCGCAAGTTGGTGGTGGTTATTAATCA 764
Db 633 CAATACGGGATTAGAGCGCTGTATGGGACCGGATTTCTAGAGATTTGATTAAGATATATCA 692
QY 765 ATTCGCTAGAGATCTAACGTTTAGGGGTATAGTCTAGTGGCACTATTCCCAAGCTATGA 824
Db 693 ATTTAGAGAGAAATTAACACTAATCTGTATAGATATCTGTTCTCTATTTCCGAACTATGA 752
QY 825 CACTCGCACTTATCCAAATAAATAACGAGTGTCTCAGTTTAAACAAGGGAAGTTTATACAGCCG 884
Db 753 TAGTAGAACGTATCCAAATTCGAACAGTTTCCCAATTTAAACAAGAAATTTATCAAAACCC 812
QY 885 A--ATTGGAGCAACAGGGGTAAATATGCGCAAGTATGAATTTGGTATTAATAATCACT 942
Db 813 AGTATTAGAAATTTTGTAGTGTGTTTTCGAGGCTCGGCTCAGGGCATAGAAAGATAT 872
QY 943 TCGTTTTCGCTATAGAGACTCGGGTTATCCGAAGCCGCACTACTCTGTTGTTTCTAGAA 1002
Db 873 TAGGAGTCCACATTTGATGGATATCTTAAACAGTATTAACCATCTATACGATGCTCATAG 932
QY 1003 CAACTTACAATTTTATAGCACTTCATCAGATGGAGTGTCTACTAGGCATATGACTTACTGG 1062
```


QY	3216	CTGTCGGGGCGTGCTATATCCTCCGTGTCAACAGCGGTACAAACAGAGGATATCGAGAGG	3275
Db	3059	CTGTCGGGGTGGTGGCTATATCCTTCGTGTCAACAGCGTACNAGGAGGGATATGAGAGG	3128
QY	3276	TTGTGTTACGATCCATGAAATCGAGAACATATACAGAGAACTAAAAATTTAAAACTGTGA	3335
Db	3129	TTGGTTAACCAATCATGAGATCGAGAACATATACAGAGCAACTGAAGTTTAGCAACTGCGT	3188
QY	3336	AGAAGAGGAAGTGTATCCAAACGGATACAGGAACCTGTAAGTTATATCTCCACACCAAG	3395
Db	3189	AGAAGAGGAATCTATCCAAATACACGGTTAACGTGTAAATGTTATCTGTAATCAAGA	3248
QY	3396	TACAGCAGTATGTAATTTCCCGTAATGCTGGATATGAGGATGCAATATGAAGTTGATACTAC	3455
Db	3249	AGAAATACGGAGTCCGT-----ACACTTCTCGTAAATCGAGGATATAACGMACTCCTTC	3302
QY	3456	AGCATCTGTTAAATACAAAACCGACTTATGAAGAGAAACGTTATACAGATGTACGAGAGA	3515
Db	3303	CGTACCAGCTGATTTACGGTCAGTCTATGAAGAAATAATCGTTATACAGATGGACGAAGA	3362
QY	3516	TAATCATTTGTAATATGACAGAGGATATGTAATTTACACCACTACCACTGCTGGTTATAT	3575
Db	3363	GAATCTCTTGTAATTTAACAGAGGATATAGGGATTTACCGCCACTACCACTGCTGGTTATGT	3422
QY	3576	GACAAAAGAAATTAGAATACTTCCACAGAAACCGGATAAGGTATGGATTGAGATTTGGAGAAC	3635
Db	3423	GACAAAAGAAATTAGAATACTTCCAGNAACCGGATAGGTTATGGATTGAGATTTGGAGAAC	3482
QY	3636	GGAGGGAAGCTTTATGTTAGACAGCGTGGAAATTTACTCCTTATGGAGGAATAG	3687
Db	3483	GGAGGGAACATTTATCGTGGACAGCGTGGAAATTTACTCCTTATGGAGGAATAG	3534

RESULT 14

```

US-10-672-163-25
; Sequence 25, Application US/10672163
; Publication No. US20040093637A1
; GENERAL INFORMATION:
; APPLICANT: Malvar, Thomas
; APPLICANT: Gilmer, Amy Jelen
; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad Spectrum
; TITLE OF INVENTION: delta-Endotoxins
; FILE REFERENCE: 11792.0215.DVUS02 MECO:215--2
; CURRENT APPLICATION NUMBER: US/10/672,163
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: US 09/997,914
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/361,040
; PRIOR FILING DATE: 1999-03-02
; PRIOR APPLICATION NUMBER: US 08/754,490
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3531)
US-10-672-163-25

```

QY	345	TGCTAGGAATACGGCACTTCTGCTCGATTACAAGGTTTAGGAGATTCTTTAGAGCCCTATCA	404
Db	273	CGCTAGGAACCAAGCCATTCTAGATTAGAAGACTTAAGCAATCTTTATCAAAATTTACGC	332
QY	405	ACAGTCACATTGAAGATTGGCTAGAAACCGTATGATGCAAGAACGAGAAGTGTTCTTTTA	464
Db	333	AGAACTTTTAGAGAGTGGGAAGCAGATCCTACTAATCCAGATTTAAGAGAAGAGATGCG	392
QY	465	TACCCAAATATATAGCCCTTAGAACTTGTATTTCTTAATGCGATGCGCGCTTTTCGCAATTAG	524
Db	393	TATTCAATTCAAATGACATGAACAGTGCCTTACAACCGCTATTCTCTCTTTTGGCAGTTCA	452
QY	525	AAACCAAGAGTTCCATTATTAATCGTATATGCTCAAGTGCAGAAATTTACACCTATTATT	584
Db	453	AAATTATCAAGTTCCTCTTTATATCAGTATATGTTCAAGCTGCAAAATTTACATTTATCAGT	512
QY	585	ATTGAGAGATGCTCTCTCTTTTGGTAGTGAATTTTCGGCTTACATCGCAGGAAATTCACG	644
Db	513	TTTGGAGAGTGTTCAGTCTTTGGCAAAAGTGGGGATTTGATGCGCGCATATCAATAG	572
QY	645	TTATTATGAGCCCAAGTGGAAACAAACGAGAGATTATTCCGACTATTGCGTAGAATGTA	704
Db	573	TCGTTATAATGATTTAACTAGAGCTTATTGGCAACTATACAGATCATGCTGTACGCTGTA	632
QY	705	TAATACAGCTTAAATAGCTTGAGAGGGACAAATGCGCAAGTTGGGTGCGTTAATAATCA	764
Db	633	CAATACGGGATTAGAGCGTGTATGGGGACCGGATTTCTAGAGATTGGATAAGATATAATCA	692
QY	765	ATTCGTTAGAGATCAACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGA	824
Db	693	ATTTAGAAGAGAAATTAACACTAACTGATTAGATATCGTTTCTTCTATTTCGGACTATGA	752
QY	825	CACCTCGCACTTATCCATAAATACGAGTGTCTCAGTTAAAGGGAAGTTTATACAGAGCC	884
Db	753	TAGTAGAACGTTCCAAATTCGAAACAGTTTCCCAATTAACAGAGAAATTTATACAAACC	812
QY	885	A--ATTGGAGCAACAGGGGTAAATATGGCAAGTATGAATTTGGTATAATAATATGCACCT	942
Db	813	AGTATTAGAAAAATTTGATGGTAGTTTTCGAGGCTCGGCTCAGGGCATAGAAAGAATAT	872
QY	943	TCGTTTTCGCTATAGAGACTGCGGTTATCGAAGCCCGCATCTACTTGTATTTTCTAGAA	1002
Db	873	TAGAGTCCAAATTTGATGGATATPACTTAACAGPATTAACCATCTATACGGATGCTCATAG	932
QY	1003	CAACTTACAAATTTTAGCACTTTCATACGATGGAGTGTCTACTAGGCATATGACTTACTGG	1062
Db	933	GGGTTATTATTGGTCAGGCAATCA-----AATAATGGCTTCTCTGTAGGGTT	983
QY	1063	CGGGGGCACAAATTCATCTCGGCCAATTAGAGGCGGATTAATACTCAACGCAATGGG	1122
Db	984	TTGCGGGCCAGAAATTCATCTTTCGCTATATGGAACCTATATGGGAA-----ATGCA	1032
QY	1123	TCTACCAATACTCTTATTAATCTGTAAAGATTATCATTTCTCTCGAGAGCTATATTGG	1182
Db	1033	GCTCCACACAAACGATTTGTCTCACTAGGTCAGGGCGGTATAGAACATTTCTGTC	1092
QY	1183	ACTGAATCATATGACGAGTGCTTCTATGGGAAATTTACCTTGAACCTATTCATGTGTGTC	1242
Db	1093	ACTTTATAGAAGACCTTTTAATATAGGATAAAATAATCAACAACTATCTGTTCTTGAC	1152
QY	1243	CCTACTGTTAGATTTAATTTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAAC	1302
Db	1153	GGGAC-----AGAAATTTGCTTATGGAACCTCTCAAAATTTGGCCATCGCTGTATACAGAA	1208
QY	1303	TATAGTCAACCCCTATGAGTCACCTGGGCTTCAATTTAAAGATTTCAGAAACTG--AATTACC	1361
Db	1209	AAGCGGAACGGTAGATTTCGCTGGATGAATACCGCCACAGAAATAACAACTGCACTAG	1268
QY	1362	ACCAGAAACAAAGACGACCAAAATTAAGATCATATAGTCATAGGTTATCTCACATAGG	1421
Db	1269	GCAAGGATTTAGTCATCGAATTAAGCCATGTTTCAATGTTTTCGT-----TCAGGCTTTAG	1322
QY	1422	GCTCATTTCAAACTATAGGGTGATGTACAGATATATTTCTTGAGCGCAACGGTAGTGCAGA	1481

Db	1323	TAATAGTAGTGAAGTATAAAGAGCTCCAATGTTTTCTTGGACGACCGTAGTGCAAC	1382	Db	2349	ATGGCGCTTTTCAGGCCAAAGTCCAATCGAAAGTGTGGAGAGCCGAATCGATGCGCGC	2408
Qy	1482	TGCTACAAATACCATTTAGTTCAGATAGCATACACAAATACCATTTGGTAAATCATTTCA	1541	Qy	2559	ACATTTTCAATGGAAATCTGTAGTCTAGATTTCTTCTGCGAGAGATGGAGAAAATTTGCGCA	2618
Db	1383	CCCTACAAATACAAATTTGATCCGGAGAGGATTAATCAATACCATTTGGTAAAGGACATAC	1442	Db	2409	ACACCTTCAATGGAAATCTGTAGTCTAGATTTCTTCTGTTAGGGATGGAGAAAAGTGTGCCCA	2468
Qy	1542	CCTTAATTCAGGTACCTCTGTAGTTCAGTGGCCAGGATTTACAGAGGGGATATAATCCG	1601	Qy	2619	TCATTTCCCATCATTTCTCTTTGGATATTGATTTGGATCGACAGACTTGTGCATGAGAACT	2678
Db	1443	ACTTCAGTCAGGTACTACTGTTGTAAGAGGCCCGGGTTTACGGGAGGATATTCTTCG	1502	Db	2469	TCATTTCCCATCATTTCTCTTTGGATATTGATTTGGATCGACAGACTTGTGCATGAGAACT	2528
Qy	1602	AACTAACGTTAATGGTAGTGTACTAAGTATGGGTCTTAATTTAATAATACATCATTA	1661	Qy	2679	AGGCGTGGGTGGTATTCAAGATTAAAGCGGAGAGGTCTATGCAAGACTAGGGAATCT	2738
Db	1503	ACGAACAAGTGGAGGACCAATTTGCTTATATCTATTTGTTAATAATAAATGGCAATTA	1562	Db	2529	AGGTGTATGGGTGATCTTTAAAGATTAAAGCGCAAGATGGGACGCAAGACTAGGGAATCT	2588
Qy	1662	GGGTATCGCGTGAAGGTTGCTTATGCTGCTCTCAAAATGCTCTGAGGAGGTTAACTGT	1721	Qy	2739	GGATTTTATTCGAAGAGAAACCATTTATAGGAGAGCACTGTCTGCTGAGAGAGCAGA	2798
Db	1563	AMGTATCTGCAAGATACGCTATGCTCTCAATCAATCTAAGAAATTTACGTAACGGT	1622	Db	2589	AGAGTTTCTCGAAGAGAAACCATTTAGTAGGAGAGCGCTAGCTGTGTGAAAAGAGCGGA	2648
Qy	1722	CGGAGGAGTACTACTTTTGATCAAGGATTCCTAGTACTATGAGTGCATAATGAGTCTTT	1781	Qy	2799	GAATAAATGGAGAGACAAACCGTGAATACTCAATTTGAAAACAAACCGAGTATATACAGA	2858
Db	1623	TGAGGTGAACGGATTTTGTGCTGCTCAATTTAAACAAACCAATGGATACGGTGACCCATT	1682	Db	2649	GAATAAATGGAGAGACAAACCGTGAATACTCAATTTGAAAACAAATATATCGTTTATAAGA	2708
Qy	1782	GACATCTCAATCATTTAGATTTGCAAGATTTCTGTAGGTATTTAGTGCATCTGGCAGTCA	1841	Qy	2859	GGCAAAAGAGCTGTGGATGCTTTATTTGTAGATTTCTCAATATAATAGATTAACAAGCGGA	2918
Db	1683	AACATTTCCAATCTTTTAGTTACGCAACTATTAAATACAGCTTTTACATTTCCCAATGAGCCA	1742	Db	2709	GGCAAAAGAGCTGTGGATGCTTTATTTGTAGATTTCTCAATATAATAGATTAACAAGCGGA	2768
Qy	1842	AACTGCT--GGAATAAGTATAAGTAATAATCACTAGGTAGACAAACGTTTCACTTTGATAA	1898	Qy	2919	TACAAACATTCGCGATGATTTCATGCGGAGATTAATCTGTTTCAATCGAAATTCGAGAGCTTA	2978
Db	1743	GAGTAGTTTTCAGTAGGTGCTGATCTTTTGTAGTTTTCAGGGAATGAATTTATATAGACAG	1802	Db	2769	TACGAATATTGCCATGATTTCATGCGGAGATTAATCTGTTTCAATCGAAATTCGAGAGCTTA	2828
Qy	1899	AAATGAATTCATTTCAATTTACTGCAACCTTTGAAAGCAGAAATACGATTTAGAAAGGCGCA	1958	Qy	2979	TCTGTCAAGAAATATCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTAAGAAG	3038
Db	1803	ATTGAAATTTGATTTCCAGTTACTGCAACATTTGAAGCAGAAATATGATTTAGAAAGGACCA	1862	Db	2829	TCTGCTGAGCTGTCTGTGATTTCCGGGTGTCAATGCGGCTATTTTGAAGAAATTAAGAAG	2888
Qy	1959	AGAGCGGTGAATGCTCTGTTTACTTAATACGAATCCAAAGAAATTTGAAAACAGATGTGAC	2018	Qy	3039	TCGCAATTAATCTGCAATCTCCCTATACGATGCGGAGAAATGTGTTAAATAATGTTGATTT	3098
Db	1863	AAAGCGGTGAATGCGCTGTTTACTTCTATAAACCAATAGGGAATAAACAACAGATGTGAC	1922	Db	2889	CGGTATTTTCACTGCTATCTCTATATGATGCGGAAATGTCTATTAATAATGTTGATTT	2948
Qy	2019	AGATTTATCATATTTGATCAAGTATCAATTTAGTGGCGTGTATTCGGATGAATTTGCTT	2078	Qy	3099	TAATAATGGAATGATGCTGGAATGTAAAGGCGATGATGT--ACACAGAGCCA	3155
Db	1923	GGATTTATCATATTTGATCAAGTATCAATTTAGTGGATTTGTTTATCAGATGAATTTGTCT	1982	Db	2949	TAATAATGGAATGATGCTGGAATGTAAAGGCGATGATGTAGATGTAGAGAAACAAACAA	3008
Qy	2079	AGATGAAGAGAGAAATTAATGAGAAAGTGAATATGCGAAACGACTCAGTGATGAAG	2138	Qy	3156	TCACCGTTCTGCTCTGTTATCCAGAAATGGAGAGCAGAGTGTCAAGCAGATTTGCGGT	3215
Db	1983	GGATGAAGAGAGAAATTTGTCGAGAAAGTCAACATGCGAAGCGACTCAGTGATGAGCG	2042	Db	3009	CCACGTTGCGTCTGTTGTTTCCGGAATGGAGAGCAGAGTGTCAAGAGATTTGCTGT	3068
Qy	2139	AACTTACTTCAAGATCCAACTTCAATCCATTAATGAAGCAACAGACTTCATATCTAC	2198	Qy	3216	CTGTCGGGGCGTGGCTATATCTCCGCTGTCAAGCGTCAAAAGAGGATATGGAGAGG	3275
Db	2043	GAAATTTACTTCAAGATCCAACTTCAAGGCGATCAATAGGCAACTAGAC-----	2091	Db	3069	CTGTCGGGGCGTGGCTATATCTCCGCTGTCAAGCGTCAAAAGAGGATATGGAGAGG	3128
Qy	2199	TAATGAGCAATCGAATTTTCACTATCTCATGAACAATCTGNAACATGATGTTGGGGAAG	2258	Qy	3276	TTGTTAAGCATCCATGAAATCGAGAAACAAATACAGAGCAACTTAATAATTTAAACCTGTGA	3335
Db	2092	-----CGTGGTTGGAGAGGAAG	2108	Db	3129	TTGCGTAACCAATCATGAGATCGAGAAACAAATACAGAGCAACTGAAGTTTAGCACTGCGT	3188
Qy	2259	TGAGAAACATTAATCCAGAGAGAAATGACGATTTTAAAGAGAAATTAAGTACACTACC	2318	Qy	3336	AGAAAGAGAAATGATTTCCAGCGGATACAGGAAACCGTGAATGATTTATCTGCACACCAAG	3395
Db	2109	TACGATATTACATCCAAAGAGGAGATGACGATTTCAAGAAATTTATGTCACACTACC	2168	Db	3189	AGAAAGAGAAATCTATCCAAATACACCGTAACTGTTAATGATTTATCTGTAAATCAAGA	3248
Qy	2319	GGGGAATTTTATGATGTTATCCGACGTTATTTATATCAAAAAATAGGAGAGTCCGAAT	2378	Qy	3396	TACAGAGTATGTTAATTTCCCGTAAATGCTGGATATGAGGATGCATATGAAGTTGATCTAC	3455
Db	2169	AGGTACCTTTGATGAGTGTATCCCAACATATTTGTTATCAAAAAATCGATGAATCAAAAT	2228	Db	3249	AGAAATCGGAGGTGCGT-----ACACTTCTCGTAAATCGAGGATATAACGAAGCTCTTTC	3302
Qy	2379	AAAGCTTACTCTCTACCAATTAAGGGTATATTTGAAGTATGATCAAGATTTAGAT	2438	Qy	3456	AGCATCTTTAATTTACAAACCGACTTATGAAGAGAAACGATATACAGATGTACGAGAGA	3515
Db	2229	AAAAGCTTTTACCGTTTATCAATTTAAGAGGGTATATCGAAGATAGTCAAGACTTTAGAAAT	2288	Db	3303	CGTACAGCTGATTTATGCGTCACTATGAAGAAATTCGTATACAGATGGACGAGAGA	3362
Qy	2439	ATATTGATTCGTTAATTCGAAACATGAACATTTGGATTTCCAGGTACCGAGTCCGT	2498	Qy	3516	TAATCATTTGAAATGACAGAGGTATGTCAATTTATCCACCACTTACAGCTGGTGTATAT	3575
Db	2289	CTATTATTTGCTACATGCAAAACATGAACAGTAAATGTGCCAGGTACGGTTCCTT	2348	Db	3363	GAATCTTTGGAATTTAAGAGGGGTATAGGGATTTACGCGCACTACCGTTGGTATGT	3422
Qy	2499	ATGGCGCTTTTCAAGTTGAAGCCCAATCGGAGGTGCGGAGAACCGAATTCGATGCGACC	2558	Qy	3576	GACAAAAGAAATTAAGATACTTTCCAGAAACCGATAAGGTATGGATTTGAGATTGGAGAAAC	3635
Db				Db	3423	GACAAAAGAAATTAAGATACTTTCCAGAAACCGATAAGGTATGGATTTGAGATTGGAGAAAC	3482

```
QY 3636 GGAAGGGAAGCTTTATTGTAGACAGCGTGAATTTACTCTTATTGGAGGAATAG 3687
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3483 GGAAGGAACATTTATCGTGACAGCGTGAATTTACTCTTATTGGAGGAATAG 3534
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-10-739-482-25
; Sequence 25, Application US/10739482
; Publication No. US20040132975A1
; GENERAL INFORMATION:
; APPLICANT: Mohan, Thomas
; APPLICANT: Malvar, Thomas
; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad Spectrum
; TITLE OF INVENTION: Delta-Endotoxins
; FILE REFERENCE: MECO:220--1
; CURRENT APPLICATION NUMBER: US/10/739,482
; PRIOR FILING DATE: 2003-12-18
; PRIOR APPLICATION NUMBER: US 09/636,746
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 6,242,241
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: US 6,110,464
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: US 6,017,534
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3531)
US-10-739-482-25

Query Match 38.0%; Score 1402; DB 17; Length 3534;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

QY 285 TCAGTGGGAATTTCTTAGAACATGTCGAACACTTTATAAATCAACAATAACAGAAA 344
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 TCAATGGGACGCAATTTCTGTACAAATTTGAACAGTTAATTAACCAAGATAGAAAT 272
QY 345 TGCTAGGAATACGGCACTTCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTATCA 404
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 CGCTAGGAACGAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTATCAAATTTACGC 332
QY 405 ACAGTCACTTGAAGATTGGCTAGAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTA 464
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
333 AGAATCTTTTAGAGAGTGGGAAGCAGATCTTACTAATCCAGCATTTAAGAGAAGAGATGCG 392
QY 465 TACCANTATATAGCCTTAGAATCTTGATTTCTTAATGCGATGCCCTTTTCGCAATTAG 524
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
393 TATTCATTCATAGATGACATGAACAGTGCCTTTACAACCGCTATTCTCTCTTTTTCAGTTCA 452
QY 525 AAACCAAGAAGTTCCTATTATTAATGTTATGTTATGTTATGCGATGCCCTTTTCGCAATTAG 584
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
453 AAATTATCAAGTTCCTCTTTATCATGATATGTTCAAGCTGCAAAATTTATCATTTATCAGT 512
QY 585 ATTGAGAGATGCCCTCTCTTTTGGTATGTAATTTGGGCTTTACATGCGAGGAATTCACAG 644
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
513 TTTGAGAGATGTTTCAGTGTGTTGGACAAAGGTGGGATTTGATGCGGCACTATCAATAG 572
QY 645 TTATTATGAGCGCAAGTGGACAAACAGAGAGATTTATCCGACTATTCCGTTAGAAATGTA 704
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
573 TCGTTATATGATTAACTTAGGCTTATTGGCAACTATACAGATCATGCTGTACGCTGTA 632
QY 705 TAATACAGGTCTAAATAGCTTGAGAGGGGCAAAATGCCGCAAGTTGGGTGCGTTAATATCA 764
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
633 CAATACCGGATTTAGAGCTGTATGGGGACCGGATTTCTAGAGATTGGATAAGATAATCA 692
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
765 ATTCCTGAGAGATCTTAACGTTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGA 824
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
693 ATTTAGAAGAGAATTAACACTACTGTATTAGATATCGTTTCTATTTCGGAACATATGA 752
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
825 CACTCGCACTTATCCAAATAAATACGAGTGTCTAGTTAAACAAGGAAGTTTATACAGACGC 884
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
753 TAGTAGAACGTATCCAAATTCGAACAGTTTCCCAATTAACAAGAGAAAATTTATACAAACCC 812
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
885 A--ATTGGAGCAACAGCGGTAAATATAGCAAGTATGAATTTGGTATATATATATATGACCT 942
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
813 AGTATTAGAAAATTTTGTAGTGTAGTTTTCGAGGCTCGGTCTAGGGCATAGAAAGATAT 872
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
943 TCGTTTTTCGCTATAGAGACTGCGGTTATCCGAAGCCCGCATCTACTTCTGATTTCTAGAA 1002
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
873 TAGGAGTCCACATTTGATGGATATATTAACAGTATAACCACTATACGATGCTCATAG 932
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1003 CAACTTACAATTTTATAGCACTTCATCAGATGGAGTGTCTACTAGGCATATAGTACTG 1062
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
933 GGGTTATTATTATGTGTCAGGGCATCA-----AATAATGGGCTTCTCTGTAGGGTT 983
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1063 CGGGGGCAACAATTCATCTCGGCAATAGAGGGGATTAATAACCTCAACGATGGG 1122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
984 TTGCGGGCCAGAAATTCATTTTCGCTATATGGAATATGGGAA-----ATGCA 1032
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1123 TCTACCAATCTCTATTAACTCTGTAGATATCAITCTCTCTCGAGAGATATATG 1182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1033 GTCCACAACAACGATTGTGCTCAACTAGGTCAAGGCGGTATAGAACATATATCGTCC 1092
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1183 ACTGAATCATATGACAGAGTGTCTATGCGGGAATTTTACCTTGAACCTTATCATGGTCTC 1242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1093 ACTTTATATAGAGACCTTTTAAATATAGGGATTAATATCAACACTATCTGTTCTTGAC 1152
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1243 CCTACTGTAGATTAAATTTTAGGAACCTCAGAAATCTTTTGAAGAGGTAAGTCTAAC 1302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1153 GGGAC----AGAAATTTGCTTATGGAACCTCTCAAAATTTGCCATCGCTGTATACAGAA 1208
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1303 TATAGTCAACCTATGAGTCACTGGGCTTCAATTTAAAGATTCAGAACTG-AATTACC 1361
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1209 AAGCGGAACGTAGATTTCGCTGGATGAATACCGCCACAGAAATTAACACGTGCCACCTAG 1268
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1362 ACCAGAAACAACAGAACGCAAAATTAATGAATCATATATAGTATAGTATCTCATAGG 1421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1269 GCAAGGATTTAGTCATCGATTAGCCATGTTTCAATGTTTCGT-----TCAGSCTTTAG 1322
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1422 GCTCATTTCAAACTCTAGGGTGCATGTACCAGTATATCTTGGACGACCGTAGTGAGA 1481
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1323 TAATAGTAGTGAATATAGAGCTCAATGTTTCTTTGGACGACCGTAGTGCAAC 1382
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1482 TCGTACAAATACCAATTAAGTTACAGATAGCAATAACAATAACCAATGCTGTAATAATCAT 1541
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1383 CCTTAAATTCAGGTACCTCTGTAGTGCAGGATTTTACAGGAGGGAATATAATCCG 1442
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1542 CTTTAAATTCAGGTACCTCTGTAGTGCAGGATTTTACAGGAGGGAATATAATCCG 1601
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1443 ACTTCAGTCAAGTACTGTGTTGAAGAGGCGCGGTTTACGGGAGGAGATATCTCTCG 1502
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1602 AACTTAACGTTTAAATGGTGTACTTAAGTATGGGTCTTAATTTTAAATAATACATCAT 1661
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1503 ACGAAACAGTGGAGGACCAATTCCTTATATCTATTTGTTATATTAATTTGGGCAATTA 1562
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1662 GCGGTATCGCGTGAAGTTCGTTATGCTTCTCAACAAATGGTCTCAGGGAATATCTGT 1721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1563 AAGGTATCGTGCAGAAATACGCTATGCTCTACTACAAATCTAAGAAATTTACGTAACGGT 1622
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1722 CGGAGGAGTACTATCTTTTGTATCAAGATTTCCCTAGTACTATGAGTGCAAAATGAGTCT 1781
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1623 TGCAGGTGAACGATTTTGTGCTCAATTTAACAACAAATGATACCGGTGACCAAT 1682
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1782 GACATCTCAATCATTTAGATTTGCAAAATTTCTCTGTAGGTATTTAGTGCATCTCTGGCAGTCA 1841
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Db 1683 AACATCCAAATCTTTTGTAGTTACGCAACTATTAAATACAGCTTTTACATCCCAATGAGCCA 1742
Qy 1842 AACTGCT---GGAATAAGTATAAGTAATAATGCGAGGTAGACAAACGTTTCTACTTTGATAA 1898
Db 1743 GAGTAGTTTCACAGTAGTGTGTGATACCTTTTAGTTCAGGGAATGAGTTTATATAGACAG 1802
Qy 1899 AATTGAATTCATTCCTCAATTTACTGCAACCTTCGAGCAGAAATACGATTTAGAAAGGCGCA 1958
Db 1803 ATTTGAATTTGATTCAGTTACTGCAACATTTGAAGCAGAAATATGATTTAGAAAGAGCACA 1862
Qy 1959 AGAGCGGTGAATCTCTGTTTACTTAATACGAATCCAGAGATGGAACACAGATGTAC 2018
Db 1863 AAAGCGGTGAATGCGCTGTTTACTTCTATAAACCAATAGGGAATAAAACAGATGTAC 1922
Qy 2019 AGATTATCATATTGATCAAGTATCCAAATTTAGTGGCGTGTATTCCGATGAATCTGCTT 2078
Db 1923 GGAATTATCATATTGATCAAGTATCCAAATTTAGTGGATGTTTATCAGATGAATTTGTCT 1982
Qy 2079 AGATGAAGAAGAGAAATTTACTTGAGAAAGTGAATATCGGAAACGACTCAGTGTAGAAAG 2138
Db 1983 GGATGAAGAAGAGAAATTTGTCGAGAAAGTCAACATCGGAAGCGACTCAGTGTAGCGG 2042
Qy 2139 AAATCTACTCCAGATCCAAACTTCACATCATCAATGAAGCAACAGACTTCATATCTAC 2198
Db 2043 GAATTTTACTCAAGATCCAAACTTCAAAGGCATCAATAGGCACTAGAC----- 2091
Qy 2199 TAATGAGCAATCGAATTTTCAATCTATCCATGAACAAATCTGAACATGATGTTGGGGAAG 2258
Db 2092 -----CGTGGTGGAGGGAAG 2108
Qy 2259 TGAGAACATTAATCCAGGAGGAAATGACGTATTTAAAGAGAAATACGTCACTAC 2318
Db 2109 TACGGATATTACCATCCAAAGAGGAGATGACGTATTCAAGAAATATATGTCACTACC 2168
Qy 2319 GGGGACTTTTATGAGTGTATCCGACGTATTTATATCAAAATATAGAGAGTCCGAAAT 2378
Db 2169 AGGTACCTTTGATGAGTGTATCCAAACATATTTGTATCAAAATATCGATGAATCAAAAT 2228
Qy 2379 AAAAGCTTTATCTCGCTACCAATTAAGAGGATATATGAAGATAGTCAAGATTTAGAGAT 2438
Db 2229 AAAAGCTTTTACCGTTATCAATTAAGAGGATATATCGAAGATAGTCAAGACTTTAGAAAT 2288
Qy 2439 ATATTTGATTCGTTAATATGCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGT 2498
Db 2289 CTATTTAATTCGTACAAATGCAAAACATGAACAGTAAATGTGCCAGGTACGGGTTCTT 2348
Qy 2499 ATGGCGCTTTTCAGTTGAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATGCGCAC 2558
Db 2349 ATGGCGCTTTTCAGCCCAATGTCCAATCGSAAAGTGTGGAGCGGAATCGATGCGGCC 2408
Qy 2559 ACATTTTGAATGGAATCTGTATCTAGATTTGTTCTGTGAGAGATGGAGAAATATGCGCA 2618
Db 2409 ACACCTTGAATGGAATCTGTATCTAGATTTGTTCTGTAGGATGGAGAAATATGTCGCCA 2468
Qy 2619 TCATTTCCCATTTCTCTTTTGGATTTGATTTGATGTCACAGACTTTGCAATGAGAAATCT 2678
Db 2469 TCATTTCCCATTTCTCTTTTGGATTTGATTTGATGTCACAGACTTTTAAATGAGGACCT 2528
Qy 2679 AGGCGTGGGTGGTATTCAGATTTAGACCCGAGAGGTCAATGCAAGACTAGGGAATCT 2738
Db 2529 AGGTGTATGGGTGATCTTTTGAAGTTAAGACCAAGATGGGACCGCAAGACTAGGGAATCT 2588
Qy 2739 GGAATTTTATGAAGAAACCAATTTATAGGAGAGCACTGTCTGTGTGAAGAGAGCAGA 2798
Db 2589 AGAGTTTCTCGAGAGAAACCAATTTAGTAGGAGAGCGGTAGCTGTGTGAAGAGCGGA 2648
Qy 2799 GAAAAATGGAGAGACAAACGTTGAAAACTCAATTTGGAACCAAAACGAGTATATACAGA 2858
Db 2649 GAAAAATGGAGAGACAAACGTTGAAAAATTTGGAATGGGAAACAAATATCGTTTATAGA 2708
Qy 2859 GCGAAAGAGAGCTGTGGATGCTTTTATTTGTAGATTTCTCAATATATAGATTACAGCGGA 2918
Db 2709 GCGAAAGAGATCTGTAGATGCTTTTATTTGTAACTCTCAATATGATCAATTTACAGCGGA 2768

Qy 2919 TACAAACATTTGGCATGATTTCATGCGCAGATAAACTTCTTTCATCGAAATTCGAGAGCCTTA 2978
Db 2769 TACGAATATTGCCATGATTTCATGCGCAGATAAACGTTGTTTCATGACATTCGAGAGCCTTA 2828
Qy 2979 TCTGTCAGAATATCTGTATTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTAAGAAG 3038
Db 2829 TCTGCTGAGCTGTCTGTGATTCGGGTGTCAATGCGGCTATTTTGAAGAAATTAAGAAG 2888
Qy 3039 TCGCATTTATCACTCCAACTCCCTATACGATGCGAGAAATCTCGTTAAATATGTTGATTT 3098
Db 2889 GCGTATTTTCTACTGATTTCTCCCTATATGATGCGGAAATGTCTATTAATAATGTTGATTT 2948
Qy 3099 TAATAATGGAATAGCATGCTGGAATGTAAAAAGGCGCATGTAGATGT---ACAACAGAGCCA 3155
Db 2949 TAATAATGGCTTATCTCTGCGAACGTGAAGGCGATGTAGATGTAGAGAACAAACAA 3008
Qy 3156 TCAACGTTCTGCTCTGTTTATCCAGAAATGGGAAGCAGAAAGTGTCAAGCAGTTCCGCT 3215
Db 3009 CCAACGTTCCGCTCTGTTTGTTCGGAATGGGAAGCAGAAAGTGTCAAGAAAGTTCCGTGT 3068
Qy 3216 CTGTCGCGGCGTGGCTATATCTCTCGTGTCAAGCGGTACAAAGAGGATATGGAGGG 3275
Db 3069 CTGTCGCGGCGTGGCTATATCTCTCGTGTCAAGCGGTACAAAGAGGATATGGAGAGG 3128
Qy 3276 TTGTGTAAACGATCCATGAAATCGAGAAACAAATACAGACGAACTAAAAATTTAAAAA 3335
Db 3129 TTGCGTAACCATTCATGAGATCGAGAACAAATACAGACGAACTGAAGTTTAGCACTGCT 3188
Qy 3336 AGAAGAGAAAGTGTATCCAAAGGATACAGGAACTGTGAATGATTTACTGCAACCAAGG 3395
Db 3189 AGAAGAGAAATCTATCCAAATTAACCGGTAAACGTTAAATGATTTACTGTAAATCAAGA 3248
Qy 3396 TACAGCAGTATGTAATTTCCCGTAAATGCTGGATATGAGGATGCATATGAAGTTGATCTAC 3455
Db 3249 AGAATACGAGGTCGCT-----ACACTTCTCGTAAATCGAGGATATAACGAACTCCCTTC 3302
Qy 3456 AGCATCTGTTAAATTAACAAACCGACTTATGAAGAGAAACGTTATACAGATGTACGAAGAGA 3515
Db 3303 CGTACAGCTGATTTATGCGTCAGTCTATGAAGAAATATCGTATACAGATGGAGAGAGA 3362
Qy 3516 TAATCATTTGGAATATGACAGAGGATATGTAATTTATCCACATACAGCTGTTGTTATAT 3575
Db 3363 GAATCTTTGTAATTTAAACAGAGGATATAGGGAATTAACACGCACTTACCAAGTTGGTTATGT 3422
Qy 3576 GACAAAGAAATAGAACTTTCCAGAAACCGATAAGGTATGGATTTGAGATTGGAGAAC 3635
Db 3423 GACAAAGAAATAGAACTTTCCAGAAACCGATAAGGTATGGATTTGAGATTGGAGAAC 3482
Qy 3636 GGAAGGGAAGTTTATTTAGACAGCGTGAATTTACTCTTATGGAGGAATAG 3687
Db 3483 GGAAGGGAACATTTATCTGTCGACAGCGTGAATTTACTCTTATGGAGGAATAG 3534

Search completed: November 21, 2004, 16:50:28
Job time : 1775 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 08:54:06 ; Search time 10040 Seconds
(without alignments)
13381.802 Million cell updates/sec

Title: US-10-614-524-1
Perfect score: 3687
Sequence: 1 ttgacttcaaatagaaaa.....tactccttatggaggaatag 3687

Scoring table: IDENTITY_NUC
Gapop 10.0., Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	1.7	880	AZ669474	ENTY88TR
2	59	1.6	908	AZ548467	ENTBK30TR
3	57.4	1.6	843	AZ551618	ENTDV54TR
4	55.8	1.5	748	BX137558	Danio rer
5	55.2	1.5	931	BH160272	ENTQV49TR
6	52.8	1.4	1101	CNS0039G	AL063921 Drosophil
7	52.4	1.4	877	AZ531291	ENTBQ34TR
8	52	1.4	849	AZ546009	ENTFW53TF
9	51.8	1.4	886	BX422107	EX422107
10	51.8	1.4	1101	CNS008VL	AL069706 Drosophil
11	51.2	1.4	900	AZ549980	ENTDD94TF
12	51	1.4	1101	CNS008MH	AL069378 Drosophil
13	51	1.4	1337	AG280055	Mus muscu
14	50.6	1.4	912	AZ551092	ENTFJ22TF
15	50.2	1.4	467	BH163521	EST566044
16	49.8	1.4	1253	AG289789	AG289789 Mus muscu
17	48.8	1.3	1064	BH416119	OP21205 M
18	48.6	1.3	890	AZ530768	ENTBH54TF
19	48.4	1.3	543	B1397189	kt58a11.y
20	48.4	1.3	905	AZ550256	ENTEV58TR
21	48.4	1.3	1101	CNS0106X	AL098595 Drosophil
22	47.6	1.3	828	AZ669217	ENTFV38TR
23	47.6	1.3	906	AZ529180	ENTCE59TR
24	47.6	1.3	939	AZ540454	ENTDF66TF

25	47.6	1.3	1101	9	CNS0100X	AL098379 Drosophil
26	47.4	1.3	654	4	BM164708	EST567231
27	47.4	1.3	997	9	CNS005TE	AL060757 Drosophil
28	47	1.3	811	4	BJ403374	BJ403374
29	47	1.3	976	8	BH149983	ENTQ933TF
30	46.8	1.3	531	4	BJ363051	BJ363051
31	46.8	1.3	568	4	BJ363776	BJ363776
32	46.8	1.3	582	4	BJ366397	BJ366397
33	46.8	1.3	782	9	AG444032	AG444032 Mus muscu
34	46.8	1.3	1101	9	CNS0026Z	AL097301 Drosophil
35	46.8	1.3	1307	9	AG347197	AG347197 Mus muscu
36	46.6	1.3	537	5	BU498625	BU498625 PFESTOa9
37	46.6	1.3	548	4	BI670654	BI670654 PFESTOa0
38	46.6	1.3	556	4	BI814240	BI814240 PFESTOa1
39	46.6	1.3	585	5	BU496406	BU496406 PFESTOa0
40	46.6	1.3	921	8	BH149893	BH149893 ENTPK44TR
41	46.4	1.3	680	4	BJ384484	BJ384484
42	46.4	1.3	683	4	BJ386441	BJ386441
43	46.4	1.3	826	4	BJ383101	BJ383101
44	46.4	1.3	875	8	AZ671853	ENTLQ71TF
45	46.2	1.3	609	4	BJ444006	BJ444006

ALIGNMENTS

RESULT 1
AZ669474
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ669474 880 bp DNA linear GSS 14-DEC-2000
ENTY88TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
AZ669474 GI:11806620
GSS.
Entamoeba histolytica
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
1 (bases 1 to 880)
Loftus B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@igr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library
Seq primer: M13-Reverse
Class: Shotgun
High quality sequence start: 22
High quality sequence stop: 853.
Location/Qualifiers
1..880
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHO51; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999).

ORIGIN

```
Query Match      1.7%; Score 61; DB 8; Length 880;
Best Local Similarity 46.4%; Pred. No. 0.00037;
Matches 199; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

QY 3252 GTACAAGAGGGATATGGAGAGGGTGTGTAAACGATCCCGTAATCGAAGACAAATACAGA 3311
Db 12 GTAAGAGAGAGAGATGAGATGATGATGAAGAGAGAGATGATGATGATGATGATGATGATGAT 71
QY 3312 CGAACTAAATTTAAATACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3371
Db 72 AGATGATGATGAAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 131
QY 3372 TAATGATTATACGACACCAAGGTACAGCAGTATGTAATTCCTGTAATTCCTGTAATTCCTG 3431
Db 132 AGAAGATGATGATGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
QY 3432 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3491
Db 192 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 251
QY 3492 AACGTATACAGATGTACGAGAGAGATTAATCTGTAATTCCTGTAATTCCTGTAATTCCTGTA 3551
Db 252 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 311
QY 3552 TCACCACTACAGCTGGTGTATATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3611
Db 312 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371
QY 3612 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3671
Db 372 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
QY 3672 CCTTATGGA 3680
Db 432 CTATGATGA 440
```

```
RESULT 2
LOCUS      AZ548467
DEFINITION ENTX30TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, genomic survey sequence.
ACCESSION  AZ548467
VERSION    AZ548467.1
KEYWORDS   GSS.
SOURCE     Entamoeba histolytica
           Entamoeba histolytica
           Eukaryota; Entamoebidae; Entamoeba.
REFERENCE  1 (bases 1 to 908)
AUTHORS   Loftus, B., Van Aken, S. and Fraser, C.
TITLE     Determination of clone end and sequences from Entamoeba histolytica
           HM1:IMSS sheared DNA library
JOURNAL    Unpublished (2000)
COMMENT    Contact: Brendan J Loftus
           Department of Eukaryotic Genomics
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0208
           Fax: 301 838 3543
           Email: bjlloftus@tigr.org
           Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
           DNA library
           Seq primer: M13-Reverse
           Class: shotgun
           High quality sequence start: 17
           High quality sequence stop: 828.
           Location/Qualifiers
             1..908
               /organism="Entamoeba histolytica"
```

```
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/institute="Vector: PHOS1; Site 1: Bat 1; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark, a
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
```

ORIGIN

```
Query Match      1.6%; Score 59; DB 8; Length 908;
Best Local Similarity 44.2%; Pred. No. 0.0012;
Matches 337; Conservative 0; Mismatches 420; Indels 6; Gaps 2;

QY 2793 AGCAGAGAAAAAATCGAGAGACAAACGTCGAAAAAATCTACAAATTCGAAACCAAAACGAGTATA 2852
Db 66 AGAAGATGATGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 125
QY 2853 TACAGAGCAAAAGAGCTGTGGATCTTTATTTTGTAGATCTCTCAATATAATAGATTACA 2912
Db 126 AGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 185
QY 2913 ACGGATACAAACATTTGGCATGTTTCATCGGCAGATTAATCTGTTTCATCGAATTCGAGA 2972
Db 186 AGATGATG---ATGATGAAGAGAGAGATGATGATGAAGATGATGAAGACGATGAAGA 242
QY 2973 GGCCTTATCTGTGAGAAATATCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAAT 3032
Db 243 CGACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
QY 3033 AGAAGTCGCATATACCTGCAATCTCCCTATACGATGCGAGAAATGTCGTTHAAATGG 3092
Db 303 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
QY 3093 TCATTTTAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3152
Db 363 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
QY 3153 CCATCACCCCTCTGCTCTTATCCGAAATGGGAAGCAGAGTGTCTCAACAGCAGTTCCG 3212
Db 423 CGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 3213 CGTCTGTCGGGGCGGTGCTATATCTCCGTGTCAAGAGATGATGATGATGATGATGATGATGATGATGAT 3272
Db 483 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
QY 3273 GGGTGTGTAACGATCCATGAAATCGAAGCA---ATACAGACGAACTAAATTTTAAAAA 3329
Db 543 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
QY 3330 CTGTGAAGAGAGAGAGTGTATCCACCGGATACAGGAACGTCGTAAATGATGATGATGATGATGATGAT 3389
Db 603 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
QY 3390 CCAGGTACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3449
Db 663 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
QY 3450 TACTACAGCATCTGTTAAATTTACAAACCGACTTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3509
Db 723 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
QY 3510 AAGAGATAATCATTTGTGAATATGACAGAGCGGTATGTAATTTAT 3552
```

```

Db      783 AGAAGTGAAGATGATGAAGATGATGAAGATGATGAAGATGAT 825

RESULT 3
LOCUS   AZ551618/c
DEFINITION Entamoeba histolytica 843 bp DNA linear GSS 14-NOV-2000
          genomic, histolytica Sheared DNA Entamoeba histolytica
          genomic, genomic survey sequence.
ACCESSION AZ551618
VERSION   1
KEYWORDS  GSS
SOURCE    Entamoeba histolytica
          Entamoeba histolytica
          Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 843)
AUTHORS   Loftus,B., Van Aken,S. and Frazer,C.
TITLE     Determination of clone end sequences from Entamoeba histolytica
JOURNAL   HMI:IMSS sheared DNA library
COMMENT   Unpublished (2000)
          Contact: Brendan J Loftus
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0208
          Fax: 301 838 3543
          Email: bjloftus@tigr.org
          Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
          DNA library
          Seq primer: M13-Reverse
          Class: shotgun
          High quality sequence start: 39
          High quality sequence stop: 838.
          Location/Qualifiers
            1..843
              /organism="Entamoeba histolytica"
              /mol_type="genomic DNA"
              /strain="HMI:IMSS"
              /db_xref="taxon:5759"
              /clone_lib="Entamoeba histolytica Sheared DNA"
              /note="Vector: pPOS1; Site 1: Bat I; Constructed at The
              Institute for Genomic Research (TIGR), Rockville, MD.
              Genomic DNA isolated from broth cultures of E. histolytica
              using a method described by Clark and Diamond (Clark,
              C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
              method for isolate identification. Exp. Parasitol.
              77:450.). The DNA was mechanically sheared to give a
              tight size distribution (~2 kb). The v + i method used for
              the library construction is described in detail in Smith,
              H.O. and Venter, J.C. (Making small insert libraries for
              whole genome shotgun sequencing projects. In Genome
              Sequencing: A Practical Approach, eds. M. Vaudin and B.
              Barrell, Oxford University Press, 1999)."
```

Query Match 1.6%; Score 57.4; DB 8; Length 843;
 Best Local Similarity 43.8%; Pred. No. 0.0031;
 Matches 348; Conservative 0; Mismatches 441; Indels 6; Gaps 2;

```

QY      2793 ACCGAGAAATGCGAGACCAACGTCGAAAATCTACAAATGGAACAAACGAGTATA 2852
          |||||
Db      812 AGAAGATGATGAAGATGATGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 753
          |||||
QY      2853 TACAGAGGCAAGAGCGTCGATGCTTTATTTCTAGATTTCTCAATATATAGATTACA 2912
          |||||
Db      752 AGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGA 693
          |||||
QY      2913 AGCGGATCAAAATGCGCATGATTCATCGCGCAGATAAATGTTTCATCGGAATTCGAGA 2972
          |||||
Db      692 AGATGAT---GATGATGAAGAGAGAGAGATGATGATGAAGATGATGAAGATGATGAAG 636
          |||||
QY      2973 GCGTTATCTGTCAGATTTATCTGTTATCCCGGGTGAATCGCGGAATTTTGAAGATT 3032
          |||||
Db      635 CGACGAATATGAATTTAGAAGATGATGATGATGAAGAGAGATGATGATGAAGAGAGA 576
          |||||

FEATURES
source
  BXL37558
  BXL37558
  DEFINITION Dario rerio genomic clone DKEY-95C15, genomic survey sequence.
  ACCESSION BXL37558
  VERSION    BXL37558.1
  KEYWORDS   GSS.
  SOURCE     Dario rerio (zebrafish)
  ORGANISM   Dario rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 748)
AUTHORS     Humphray,S.J., Huckle,E. and Durham,J.L.
TITLE       Direct Submission
JOURNAL     Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humphray@sanger.ac.uk Unpublished
            This sequence was generated from the SP6 end of BAC 95C15. 95C15 is
            part of the Daniokey BAC library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
            Location/Qualifiers
              1..748
                /organism="Dario rerio"
                /mol_type="genomic DNA"
                /db_xref="taxon:7955"
                /clone="DKEY-95C15"
                /tissue_type="Testis"
                /note="vector pIndigoBAC-536"

ORIGIN
  3033 AGAAGTGCATTTACTGCAATCTCCCTATACGATCGCAAAATGTCGTTAAAAATGG 3092
          |||||
  Db      575 AGAAGATGATGAAGATGATGAAGAGAGATGAAGAGATGATGAAGAGATGA 516
          |||||
  QY      3093 TGAATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3152
          |||||
  Db      515 TGAAGATGATGATGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 456
          |||||
  QY      3153 CCATCACCGTTCTGTCCTTTATCCAGATGCGAAGCGAGAGTGTCAACAGCAGTTCCG 3212
          |||||
  Db      455 CGAATATGAATTTAGAAGATGATGAAGAGAGAGATGATGATGAAGAGAGAGAGA 396
          |||||
  QY      3213 CGTCTGTCGGCGCGCTGCTATATCTCCGTCCTACAGCGTCAACAAGAGGGGATATGGAGA 3272
          |||||
  Db      395 TGATGATGAAGATGATGATGAAGAGAGATGATGATGATGAAGAGAGATGATGAAGA 336
          |||||
  QY      3273 GGGTTGTGTACGATCCATGAAATCGAGA---ACAATACAGACGAACTTAAATTTTAAAAA 3329
          |||||
  Db      335 TGATGATGAAGAGAGATGATGATGAAGAGAGAGATGATGAAGAGAGAGATGATGAAGA 276
          |||||
  QY      3330 CTGTGAAGAGAGAGAGTATCCAAACGATACAGGACGTCGTAATGATTATCTACTGCACA 3389
          |||||
  Db      275 AGATGATGAAGAGAGAGATGATGAAGAGAGATGATGAAGAGAGATGATGAAGAGATGATGA 216
          |||||
  QY      3390 CCAAGGTACAGCAGTATGTAATTCCTCGTAAATGCTGGATATGAGGATGCATATGAAAGTTGA 3449
          |||||
  Db      215 AGAAGATGATGAAGAGATGATGAAGAGAGATGATGAAGAGAGATGATGAAGAGATGATGA 156
          |||||
  QY      3450 TACTACAGCATCTGTTAATTACAAACCGACTTATGAAGAGAGAAAGTATACAGATGTGACG 3509
          |||||
  Db      155 TGATGATGAAGAGAGATGATGAAGAGATGATGAAGAGAGATGATGAAGAGAGATGA 96
          |||||
  QY      3510 AAGAGATAATCATGCTGAATATGACAGAGGGTGTGTAATTATCCACCCTACCGAGCTGG 3569
          |||||
  Db      95 AGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGAGAGATGAATAA 36
          |||||
  QY      3570 TTATATGACAAAAGA 3584
          |||||
  Db      35 TTAGAAGATGAAGATA 21
          |||||

RESULT 4
BXL37558
LOCUS
DEFINITION Dario rerio genomic clone DKEY-95C15, genomic survey sequence.
ACCESSION BXL37558
VERSION    BXL37558.1
KEYWORDS   GSS.
SOURCE     Dario rerio (zebrafish)
ORGANISM   Dario rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 748)
AUTHORS     Humphray,S.J., Huckle,E. and Durham,J.L.
TITLE       Direct Submission
JOURNAL     Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humphray@sanger.ac.uk Unpublished
            This sequence was generated from the SP6 end of BAC 95C15. 95C15 is
            part of the Daniokey BAC library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
            Location/Qualifiers
              1..748
                /organism="Dario rerio"
                /mol_type="genomic DNA"
                /db_xref="taxon:7955"
                /clone="DKEY-95C15"
                /tissue_type="Testis"
                /note="vector pIndigoBAC-536"

ORIGIN

```


REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 877)
Lofthus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from *Entamoeba histolytica*
HM1:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Lofthus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543

JOURNAL
COMMENT

HM1:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543

Fax: 301.856.3543
 Email: bjloftus@cigr.org
 Clones are derived from the Entamoeba histolytica HM:1MSS sheared
 DNA library
 Seq primer: M13-Reverse
 Class: shotgun
 High quality sequence start: 22
 High quality sequence stop: 829.
 Location/Qualifiers
 source
 1..877

location="Quadrant 18"
1. 877
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHO51; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaubin and B.
Barrell, Oxford University Press, 1999)."

ORIGIN

Query Match	1.4%	Score 52.4;	DB 8;	Length 877;
Best Local Similarity	48.0%;	Pred. No. 0.059;		
Matches 189;	Conservative	0;	Mismatches 196;	Indels 9;
Gaps				

[illegible]

Qy 3317 TAAATTTAAAACCTGTGAAGAGAGGAAAGTGTATCCAACGGATACAGGAACTGTATG 3376
 | | | | | | | | | | | | | | |
Db 391 ATGATGATGAAGATGATGAAGACCATGAAGACCGAATATGAATTAGAAGATGATGATG 332

QY 3377 ATTATACGCACACCAAGGTACAGCAGTATGTATATCCCGATATCGTGATATCAGGATG 3436

Dδ 331 AAGAAGAAGAATGATGTATGAGAAGAAGAAGATGATGTAAGATGATGATGATGAAGAAG 272

QY 3437 CATATGAAGTTGATCTACTACAGCATCTGTTAATTACAACCGACTTATGAAGAAGAAACGT 3496
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 AAGATGAAGATGATGATGAAGAGAAGATCAAGA-----TGATGATGAAGAAGAAG 221

[illegible]

3557 CACTACAGCTGTTATATGACAAAGAAATTAGAATACTTCCAGAAACCGATAAGGTAT 3616
QY
160 AAGATGATGAAGACGACGAATATGAATTAGAAGATGATGAAGAGATGATGATCAAGAAG 101
Db

QY 3617 GCATTGAGATTGCAGAAACGGAAGGAAGTTTAT 3650
||| ||| | | | | |
Db 100 ATGATGAAGAAGATATTCAGCAAGAAGAAGAT 67

and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila/bac.htm>.

```

location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end : TET3"

```

```

ORIGIN
Query Match          1.4%; Score 52.8; DB 9; Length 1101;
Best Local Similarity 15.7%; Pred.No. 0.05;
Matches 99; Conservative 264; Mismatches 269; Indels 0; Gaps 0;

Qy 1601 GAACAAACGTTAAGTGTAGTGTACTAAGTATGGCTCTAAATTTTAAATAACATCATTC 1650
      ||:  ::  :::::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db 1095 GDTTWRDTRKDDWDWTKWTWKNDKRRWAGDADRWAWDDGAGTWTATTTT:::WWW
      ||:  ::  :::::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

```

1095	Db	GGDTWDNRDTRKDDWDWTKWTTWKDRADRRWAGDADRWADDGAGTGTWTTATTTWWWWWWW	1036
1661	Qy	AGCGGTATCCGGTGAGAGTTCGGTATTACGTCTTCTCAACAATGGTCTCGAGGTAACGT	1720
1035	Db	ATWDTWDDKWWWATAKTDTTATWTTWTAWRADWAGDRGAGKGRDAAATDADGAGRRDG	976
1721	Qy	TCGGAGGGAGTACTACTTTTGTATCAAGGATTCCTCTAGTACTATGAGTGCAAAATGAGTCTT	1780
975	Db	GRRKDKRKRGDDDKGKGKKAATAKATKWDDDDWDDKDWGDKARDKADDDG	916
1781	Qy	TGACATCTCAATCATTTAGATTTTGCAGAAATTCCTGTAGGTATTAGTGCATCTGGCAGTC	1840
915	Db	AGDKDDGKGKADDDTDGTAKDDDDKDKDDWDKAKGTWGDATWAAATATDWWMGWADAD	856
1841	Qy	AAACTGCTGGAATAAGTATATAGTAATAATCGCAGGTAGACAAAAGTTTTCATTGATAAAA	1900
855	Db	WTTWDAAADDDWADDADWAKWDDADWAWGARTATATATATATATATATATATATATATATAT	796
1901	Qy	TTGAATTCATTCCAAATTTACTGCACCTTCGAAGCAGAAATACGATTTTAGAAGGGCGCAAG	1960

[illegible]

QY 1961 AGGCGGTGAATGCTCTGTTTACTTAATACGAATCCAGAAGATTGAAACAGATGTGCAG 2020

[illegible]

2081	ATGAAAGAGAGAATTACTTTGAGAAAGTGAATAATGCGAAACGACTCAGTGTGCAAGAA	2140
615	TTADRDWAAADTTTADKARDWAKARAWRRDRRAAARADRRWTTTGKTTTATTTT	556

[illegible]

333 ATGAGCAATCGAATTTTCACATCTATCCATGAA 2232
 2201 ATGAGCAATCGAATTTTCACATCTATCCATGAA 2232
 495 AAAAAAAAAAAATTTTTTTTTTTTAAWTAAA 464
 495 AAAAAAAAAAAATTTTTTTTTTTTAAWTAAA 464

DD
RESULT 7
AZ531291/c
404

AZ531291 877 bp DNA linear GSS 03-NOV-2000
 LOCUS ENTB034FR Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION genomic, genomic survey sequence.
 ACCESSION AZ531291

accession AZ531221
 version AZ531291.1 GI:11085838
 keywords GSS.
 source Entamoeba histolytica
 organism Entamoeba histolytica

ORGANISM
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.

RESULT 8
 A2546009/c
 LOCUS ENTFW37F Entamoeba histolytica Sheared DNA linear GSS 14-NOV-2000
 DEFINITION genomic, genomic survey sequence.
 ACCESSION A2546009
 VERSION A2546009.1 GI:11167130
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica
 ORGANISM Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 849)
 Loftus, B., Van Aken, S. and Frazer, C.
 AUTHORS Determination of clone end sequences from Entamoeba histolytica
 TITLE HM1:IMSS sheared DNA library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: b.loftus@tigr.org
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
 DNA library
 Seq primer: M13-Forward
 Class: shotgun
 High quality sequence start: 26
 High quality sequence stop: 796.
 Location/Qualifiers
 1..849
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HM1:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHSI; Site 1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + i method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barel, Oxford University Press, 1999)."

FEATURES source

Db 434 TGATGAAGATGAAGAAGATGATGATGAAGAAGACGAAGAAGATGATGATGA 435
 QY 3108 ATTACGATGCTGGAATGTAAAGGCGCATGTAGATGTACAACAGAGCCATCACGTTCTGT 3167
 Db 434 ATTTGAATTAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGATGATGATGA 375
 QY 3168 CCTTGTATCCAGATGGGAAGCAGAGTGTCAAGCAGATTTCGGTCTCTGCCGGCG 3227
 Db 374 AGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGA 315
 QY 3228 TGGCTATATCCTCGTGTACACGCTACAAGAGGATATGGAGAGGTTGTGTAACGAT 3287
 Db 314 TGATGATGATGAAGATGAAGAAGATGATGATGAATTTGAATTAGAAGATGA 255
 QY 3288 CCATGAATCGAACAATACACAGCAACTTAAATTTTAAAACTGTGAAGAAGAGCAACT 3347
 Db 254 AGAAGATGAAGATGATGAAGACGCAAGACGAGATGATGATGAATTTGAATTGA 195
 QY 3348 GTATCCACCGATACAGGACGTGTAATGATTATCTGCACACCAAGGTACAGCAGTATG 3407
 Db 194 AGATGAAGATGAAGATGAAGAAGATGAAGAAGATGATGAAGAAGAAAGAAGA 135
 QY 3408 TTAATCCCGTAATGCTGGATATGAGATGCATATGAAGTTGATCTACAGCATCTGTAA 3467
 Db 134 AGATGATGATGAAGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGAAGA 75
 QY 3468 TTACAACCGACTTATGAAGAAGA 3491
 Db 74 TGACGAAGATGTTGATGAAGATGA 51

RESULT 9
 BX422107
 LOCUS 886 bp mRNA linear EST 01-MAY-2004
 DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens CDNA
 clone CS0DH007YD10 3-PRIME, mRNA sequence.
 ACCESSION BX422107
 VERSION BX422107.2 GI:46930727
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 886)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 AUTHORS Full-length cDNA libraries and normalization
 TITLE Unpublished (2001)
 JOURNAL On May 13, 2003 this sequence version replaced gi:30651307.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 division of Invitrogen.
 This sequence belongs to sequence cluster 3628.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0DH007YD10SNP1&c=3628.r.

FEATURES source

Query Match 1.4%; Score 52; DB 8; Length 849;
 Best Local Similarity 43.3%; Pred. No. 0.075;
 Matches 296; Conservative 0; Mismatches 385; Indels 3; Gaps 1;
 QY 2811 AGACAAACGTTGAAACCTCAATTTGAAACCAACAGGTATATACAGCGCAAGAACG 2870
 Db 734 AGATGAATGTTAGAAATGGAACATGATACCGTTAGAAAGCAATGAATTAGAAGAAGA 675
 QY 2871 TGTGGATGCTTTATTTG----TAGATTCTCAATATATAGATTACAAGCGGATACAACAT 2927
 Db 674 TGATGATGATGAATTTGAAATTAGAACGCAAGATGATGATGATGATGATGAAGCAAGA 615
 QY 2928 TGGCATGATTCATCGCGCAGATAAACTGTTTCATCGATTTCAGAGGCTTATCTGTGACA 2987
 Db 614 AGATGAAGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
 QY 2988 ATTATCTGTTATCCCGGGTGTAAATTCGCGAAATTTTGAAGAAATTTAGAAAGTTCGCATTAT 3047
 Db 554 ATTAGAAGATGAAGAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGA 495
 QY 3048 CACTGCAATCTCCCTATACGATGCGGAGAAATCTCGTTAAATTTGATGATTTTAATATGG 3107

Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjloftus@tigr.org
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library
 Seq primer: M13-Forward
 Class: shotgun
 High quality sequence start: 20
 High quality sequence stop: 890.
 Location/Qualifiers
 1..900

FEATURES

source
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOS1; Site 1: Bet I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell, Oxford University Press, 1999)."

ORIGIN

Query Match
 Best Local Similarity 48.0%; Score 51.2; DB 8; Length 900;
 Matches 146; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 3290 ATGAATCGAGAACATACAGACGAACTAAATTTAAACCTGTCAGAGAGAGAGTGT 3349
 Db |||||
 QY 323 ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 382
 Db |||||
 QY 3350 ATCAACCGATACAGAACGCTGATGATGATGATGATGATGATGATGATGATG 3409
 Db |||||
 QY 383 ATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGA 442
 QY 3410 ATTCCCGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3469
 Db |||||
 QY 443 AAGAAGATGAAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 502
 QY 3470 ACACACCGACTTATGAAGAACAGCTATACAGATGATGATGATGATGATGATG 3529
 Db |||||
 QY 503 ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 562
 QY 3530 ATGACAGAGGTATGTAATATCCACCTACAGCTGATGATGATGATGATGATG 3589
 Db |||||
 QY 563 ATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGA 622
 QY 3590 AATA 3593
 Db |||||
 QY 623 AAGA 626

RESULT 12
CNS00EMH/c

LOCUS
 DEFINITION
 CNS00EMH 1101 bp DNA linear GSS 04-JUN-1999
 Drosophila melanogaster genome survey sequence T7 end of BAC:
 BACR29M06 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 AL069378
 AL069378.1 GI:4949521
 GSS
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE
AUTHORS
TITLE
JOURNAL

Ephyrdoidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPI-98 and was constructed by partial, EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the BDGP's and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR29M06"
 /clone_lib="RPI-98"
 /note="end : T7"

ORIGIN

Query Match
 Best Local Similarity 1.4%; Score 51; DB 9; Length 1101;
 Matches 102; Conservative 69; Mismatches 148; Indels 0; Gaps 0;

QY 2681 CGGTGGTGGTGTATTCAGATTAAAGACGACGAGGTCATGCAAGACTAGGGAATCTGG 2740
 Db |||||
 QY 932 GGGTWTGKAGWGRATKRDWAKWTRWTAWAATAATATATGATDAATAMWTGG 873
 QY 2741 AATTATTGAGAGAACCATTTATTAGGAGACGACTGTCTGTGTGAAGACGACGAG 2800
 Db |||||
 QY 872 AKGTATGGARRRRARAKGKTGTATGATWKAkakakakakakakakakakakakak 813
 QY 2801 AAAAATGGAGAGACAAACGCGTGAACAACTACATTCGAAACAAACGAGTATATACAG 2860
 Db |||||
 QY 812 AGAKAGAGTRGRWAGTATATKATRAARAADAGATAKAAGGDRATRKARRAKATRTAG 753
 QY 2861 CAAAAGAACTGTGGATGCTTTTATTGTAGATTCTCAATATAATAGATTACAAGCGGATA 2920
 Db |||||
 QY 752 KRWAkakakawagAGTGKAWAKATGTGWTATTTKAKWRKARKRWGTAKATDTTGRAAAA 693
 QY 2921 CAACATTTGGCATGATTCATCGGACGATAACTTTGTTTCATCGAATTCGAGAGCTTATC 2980
 Db |||||
 QY 692 AATAAAGAGATRTTRRRKWKAWAWATATAAGTGTAGARAARAAGAAAGTGTGTG 633
 QY 2981 TGTCAGAATTATCTGTTAT 2999
 Db |||||
 QY 632 TDGGGKXKXDTGGRWGT 614

RESULT 13
AG280055/c

LOCUS
 DEFINITION
 AG280055 1337 bp DNA linear GSS 02-JUN-2004
 Mus musculus molossinus DNA, clone:MSMg01-051D13.T7, genomic survey sequence.
 AG280055
 AG280055.1 GI:47852932
 GSS
 Mus musculus molossinus
 Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

1
REFERENCE
AUTHORS
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE
BAC end Sequences of Library MSMg01
JOURNAL
Unpublished
2 (bases 1 to 1337)
Direct Submission
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
JOURNAL
1-7-22 Sushiro-chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(S-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBACe3.6
R.Site 1 : ECORI
R.Site 2 : ECORI.
Location/Qualifiers
1. .1337
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-051D13.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
FEATURES
source

```

[illegible]

RESULT 14	AZ551092	912 bp	DNA	linear	GSS 14-NOV-2000
LOCUS	AZ551092				
DEFINITION	ENTF222F <i>Entamoeba histolytica</i> Sheared DNA				<i>Entamoeba histolytica</i>
ACCESSION	AZ551092				genomic, genomic survey sequence.
VERSION	AZ551092.1				GI:11176393
KEYWORDS	GSS.				
SOURCE	<i>Entamoeba histolytica</i>				
ORGANISM	<i>Entamoeba histolytica</i>				
	Eukaryota; Entamoebidae; Entamoeba.				

REFERENCE
1 (bases 1 to 912)
Loftus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
TITLE
HM1:IMSS sheared DNA library
JOURNAL
Unpublished (2000)
COMMENT
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjoftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 861.
Location/Qualifiers
1..912
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica sheared DNA"
/note="Vector: pHO31; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell. Oxford University Press, 1999)."

Barrell, Oxford University Press, 1999).

ORIGIN	Query Match Best Local Similarity Matches 260; Conservative	1.4%; 42.7%; 0;	Score 50.6; Pred. No. 0.17; 0;	DB 8; Mismatches 349; Indels	Length 912; 0; Gaps
Qy	3007	GTAAATCGGAATTTTGAAGAATTTAGAAGTCGCAATTATCACTGCAATCTCCCTATAC	3066		
Db	243	GAAGATGATGAAGATGATGATGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAA	302		
Qy	3067	GATCGAGAAATGTCGTTAAAATGTGTGATTTTAAATAATGGATTAGCATGCTGGAAATGTA	3126		
Db	303	GATGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAATGATGAAGAT	362		
Qy	3127	AAAGGCCATGTAGATGTACAACAGAGCCATCACCGTTCTGTCCTCTGTTATCCAGAAATGG	3186		
Db	363	GAAGATGATGAAGACCATGAAGACGACGAATATGAATTTAGAAGATGATGATGAAGAAGAA	422		
Qy	3187	GAAGCAGAAAGTGTCAACAAGCAGTTCCCGTCTGTCCCGGGCGTGGCTATATCCTCCCGTGC	3246		
Db	423	GAAGAAGATGATGATGAAGATGATGAAGAGATGAAGACGACGAATATGAAATTAGAAGAT	482		
Qy	3247	ACAGCGTACAAAGAGGATATGGAGAGGTTGTGTAACGATCCATGAATTCGAGAAACAAT	3306		
Db	483	GATGATGATGATGAAGAAGATGATGATGAAGAAGAAAGATGATGATGAAGATGAT	542		
Qy	3307	ACAGACGAACATAAATTTAAAACCTGTGAAGAGAGGAAGTGTATCCACGGATACAGGA	3366		
Db	543	GATGAAGAAGAGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGAA	602		
Qy	3367	ACGTGTAATGATTATCTACTGCACCAAGGTACAGCAGTATGTAAATTTCCCGTAAATCTGGA	3426		
Db	603	GAAGATGATGATGAAGATGATGAAGACGATGAAGACGACGAATATGAAATTAGAAGATGAT	662		
Qy	3427	TATGAGGATGCATATGAAGTTGATCTACAGCATCTGTTAAATTTTACAACCGCACTTATGAA	3486		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:33:31 ; Search time 119 Seconds
(without alignments)
3701.845 Million cell updates/sec

Title: US-10-614-524-2

Perfect score: 6479

Sequence: 1 LTSNRKNEINAIISIPAV.....ICETEGKFIVDSVEILLMEE 1228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6479	100.0	1228	4	AAB84628 Amino aci
2	6462	99.7	1228	4	Aau02039 B. thurin
3	5926.5	91.5	1227	2	Aaw44321 Bacillus
4	5926.5	91.5	1227	4	Aab19950 Bacillus
5	5912.5	91.3	1227	2	Aau02046 B. thurin
6	5740	88.6	1228	2	Aar50955 Bacillus
7	5436.5	83.9	1227	2	Aay31990 Chimeric
8	5237.5	80.8	1229	2	Aar54074 CryET5. 2
9	5237.5	80.8	1229	2	Aaw35259 Bacillus
10	5237.5	80.8	1229	2	Aaw17699 CryET5. 3
11	5237.5	80.8	1229	2	Aaw87633 CryET5 pr
12	5237.5	80.8	1229	2	Aay30923 B. thurin
13	5237.5	80.8	1229	8	Adk98479 B thuring
14	5233.5	80.8	1230	8	Adk98484 B thuring
15	5233.5	80.8	1230	8	Adk98489 B thuring
16	5233.5	80.8	1230	8	Adk98481 B thuring
17	5233.5	80.8	1230	8	Adk98491 B thuring
18	5233.5	80.8	1230	8	Adk98487 B thuring
19	5189.5	80.1	1209	4	Aau02094 Bacillus
20	5183.5	80.0	1221	4	Aau00421 B. thurin
21	5169.5	79.8	1221	4	Aau00420 B. thurin
22	5108	78.8	1186	2	Aay16796 Amino aci
23	4791	73.9	1208	4	Aau02093 Bacillus
24	4303.5	66.4	1217	4	Aau02092 Bacillus
25	3977.5	61.4	1174	2	AAR89494 B.t. toxi

26	3977.5	61.4	1174	2	AAW09043	Bacillus
27	3974.5	61.3	1174	2	AAR25825	Aar25825 Novel tox
28	3969.5	61.3	1174	2	AAR08257	Aar08257 B.thuring
29	3791	58.5	1176	2	AAW46737	Aaw46737 Amino aci
30	3791	58.5	1176	2	AAW47035	Aaw47035 Bacillus
31	3579	55.2	1170	2	AAR63231	Aar63231 Cryetal p
32	3556	54.9	1167	2	AAR54073	Aar54073 CryET4. 2
33	3556	54.9	1167	2	AAW35258	AAW35258 Bacillus
34	3556	54.9	1167	2	AAW17700	Aaw17700 CryET4. 3
35	3556	54.9	1167	2	AAW87632	Aaw87632 CryET4 pr
36	3556	54.9	1167	2	AAAY30922	B. thurin
37	3502.5	54.1	1189	2	AAW61345	Aaw61345 a mutated
38	3502.5	54.1	1189	2	AAAY17783	Aay17783 EGI2111 c
39	3502.5	54.1	1189	3	AAAY82431	Aay82431 Bacillus
40	3502.5	54.1	1189	5	AAE26471	AAE26471 Bacillus
41	3502.5	54.1	1189	5	ABG93849	ABG93849 B. thurin
42	3502.5	54.1	1189	7	ABW02350	ABW02350 B. thurin
43	3500.5	54.0	1189	2	AAW61335	Aaw61335 A mutated
44	3500.5	54.0	1189	2	AAAY17788	Aay17788 B. thurin
45	3500.5	54.0	1189	3	AAAY82396	Aay82396 Bacillus

ALIGNMENTS

RESULT 1

AAB84628
ID AAB84628 standard; protein; 1228 AA.

XX AC AAB84628;

XX XX 05-SEP-2001 (first entry)

DE Amino acid sequence of a CryIbf insecticidal protein.

XX CryIbf; insecticidal protein; CryIJD; Cry9Fa; lepidoptera; coleoptera;
KW insect pest; transgenic plant.

XX OS Bacillus thuringiensis.

XX XX WO200147952-A2.

XX XX 05-JUL-2001.

XX PF 19-DEC-2000; 2000WO-EP013184.

XX PR 28-DEC-1999; 99US-0173387P.

XX XX (AVET) AVENTIS CROPS SCIENCE NV.

XX PI Arnaut G, Boets A, Damme N, Mathieu E, Vanneste S, Van Rie J;

XX DR WPI; 2001-425619/45.

XX DR N-PSDB; AAH28240.

XX PT Novel insecticidal proteins CryIJD, Cry9Fa and CryIbf, derived from
PT Bacillus thuringiensis, useful for controlling insects in plants.

XX PS Claim 11; Page 37-41; 65pp; English.

XX CC The present sequence represents CryIbf, an insecticidal protein derived
XX CC from Bacillus thuringiensis. The specification also describes CryIJD and
XX CC Cry9Fa. The Cry proteins have activity against lepidopteran and
XX CC coleopteran insect pests. CryIbf, CryIJD and Cry9Fa polynucleotides and
XX CC polypeptides are useful for obtaining a plant with resistance to insects.
XX CC Cry polynucleotides are useful for producing transgenic plants which are
XX CC resistant to insects

XX SQ Sequence 1228 AA;

Query Match 100.0%; Score 6479; DB 4; Length 1228;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 LTSNRKNEEIIINALSIPAVSNHSTQMDLSDPARIEDSLCIAAGNNINPLVSASTVQGI 60
Db 1 LTSNRKNEEIIINALSIPAVSNHSTQMDLSDPARIEDSLCIAAGNNINPLVSASTVQGI 60

QY 61 NIAGRILGVLGVPFAGQIASFSFLVGLWPRGRDQWEIFLEHVOLINQOITENARNTA 120
Db 61 NIAGRILGVLGVPFAGQIASFSFLVGLWPRGRDQWEIFLEHVOLINQOITENARNTA 120

QY 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLTQYIALELDFLNAMPLFAIRNQEV 180
Db 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLTQYIALELDFLNAMPLFAIRNQEV 180

QY 181 LLMVYAQAANLHLLLRDASLFGSEFGLTSQIEIQRYEROVEQTRDYSDYCVWYNTGLN 240
Db 181 LLMVYAQAANLHLLLRDASLFGSEFGLTSQIEIQRYEROVEQTRDYSDYCVWYNTGLN 240

QY 241 SLRGTTAAASWRYNQFRRLDLTLGLVDLVALFPSTRTPTINTSAQLTREYVTDAGTG 300
Db 241 SLRGTTAAASWRYNQFRRLDLTLGLVDLVALFPSTRTPTINTSAQLTREYVTDAGTG 300

QY 301 VNASMNNVNNAPSFAIETAVIRSPLHLLDFLQLTIFSTSSRWSATRHMTYWRGHTIQ 360
Db 301 VNASMNNVNNAPSFAIETAVIRSPLHLLDFLQLTIFSTSSRWSATRHMTYWRGHTIQ 360

QY 361 SRPIGGGLNTHGSTNTSINPRLSPFRSDVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420
Db 361 SRPIGGGLNTHGSTNTSINPRLSPFRSDVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420

QY 421 FRNQPNTERTGANYSQPYEFGIQLKDSETELPETTERPNYESYSHRLSHIGLISQSR 480
Db 421 FRNQPNTERTGANYSQPYEFGIQLKDSETELPETTERPNYESYSHRLSHIGLISQSR 480

QY 481 VHPVYSWTHRSADRTNTISSDSTQIPLVKSFNLSGTSVSGPGTGDDIIRTNVNGS 540
Db 481 VHPVYSWTHRSADRTNTISSDSTQIPLVKSFNLSGTSVSGPGTGDDIIRTNVNGS 540

QY 541 VLSMGLNPNNTSLQRYRVRYAAASQTMVLRVTVGGSTTFDQGPSTMSANESILTSQSP 600
Db 541 VLSMGLNPNNTSLQRYRVRYAAASQTMVLRVTVGGSTTFDQGPSTMSANESILTSQSP 600

QY 601 PAEPFVGISAGSGTAGISISNNAGROTFFHDKIEFIPITATPEAEYDLERAQEAVALF 660
Db 601 PAEPFVGISAGSGTAGISISNNAGROTFFHDKIEFIPITATPEAEYDLERAQEAVALF 660

QY 661 TINTPRELKTVDVTDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQDPN 720
Db 661 TINTPRELKTVDVTDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQDPN 720

QY 721 FTSINKQPDFTISTNEQSNFTSIHQSHGHWGSENITIQEGNDVPKENYVTLPGTFNECY 780
Db 721 FTSINKQPDFTISTNEQSNFTSIHQSHGHWGSENITIQEGNDVPKENYVTLPGTFNECY 780

QY 781 PTLYLYQIGSELSKAYTRYQLRGVIEDSQDLELYLRYNAKHETLDVPGTESWPLSVES 840
Db 781 PTLYLYQIGSELSKAYTRYQLRGVIEDSQDLELYLRYNAKHETLDVPGTESWPLSVES 840

QY 841 PIGRCGPNRCAPHFENPDLDCSRDGEKCAHSHHFSIDIDIGCTDLHENLGVWVFK 900
Db 841 PIGRCGPNRCAPHFENPDLDCSRDGEKCAHSHHFSIDIDIGCTDLHENLGVWVFK 900

QY 901 IKTOEGHARLGNLEFTEEPKLLGEALSRVKRAKWRDRKREKLETKRYVTEAKEAVDA 960
Db 901 IKTOEGHARLGNLEFTEEPKLLGEALSRVKRAKWRDRKREKLETKRYVTEAKEAVDA 960

QY 961 LFVDSQYNRLQADTNIGMHAADKLVHRIREAYLSLSVPGVNAIRIPEELGRITITALS 1020
Db 961 LFVDSQYNRLQADTNIGMHAADKLVHRIREAYLSLSVPGVNAIRIPEELGRITITALS 1020

QY 1021 LYDARNVKNGDPNGLACWNVKGHDVQOSSHRSVLVTPWEAEVSQAVRVCPRGYIL 1080
Db 1021 LYDARNVKNGDPNGLACWNVKGHDVQOSSHRSVLVTPWEAEVSQAVRVCPRGYIL 1080
```

```
QY 1081 RVTAYKEGYGEGCVTIHEIENNTDELKFKNCBEEVYPTDTGTCNDYTAHQGTAVCNRN 1140
Db 1081 RVTAYKEGYGEGCVTIHEIENNTDELKFKNCBEEVYPTDTGTCNDYTAHQGTAVCNRN 1140

QY 1141 AGYEDAYEVDDTTASVNYKPTYEETVTVRRDNHCYDRGVNYPPLPAGYMTKELEYFP 1200
Db 1141 AGYEDAYEVDDTTASVNYKPTYEETVTVRRDNHCYDRGVNYPPLPAGYMTKELEYFP 1200

QY 1201 ETDKWIEIGETEGKFIVDSVELLME 1228
Db 1201 ETDKWIEIGETEGKFIVDSVELLME 1228

RESULT 2
AAU02039
ID AAU02039 standard; protein; 1228 AA.
XX
AC AAU02039;
XX
DT 29-AUG-2001 (first entry)
XX
DE B. thuringiensis toxic crystal protein, CryET64.
XX
KW Delta endotoxin; Lepidopteran-active; crystal protein; insecticide;
transgenic plant; corn; wheat; soybean; oat; cotton; rice; yre; sorghum;
sugarcane; tomato; tobacco; kapok; flax; potato; barley; turf grass;
pasture grass; berry; fruit; legume; vegetable; ornamental plant; shrub;
cactus; tree call; gypsey moth; looper; tobacco budworm; spruce budworm;
cotton leaf perforator; CryET64.
XX
OS Bacillus thuringiensis.
XX
PN WO200119859-A2.
XX
PD 22-MAR-2001.
XX
PF 13-SEP-2000; 2000WO-US025361.
XX
PR 15-SEP-1999; 99US-0153999SP.
XX
(MONS ) MONSANTO CO.
XX
PI Baum JA, Chu C, Donovan WP, Gilmer AJ, Rupar MJ;
DR WPI: 2001-281518/29.
XX
DR N-PSDB; AAS02482.
XX
PT Lepidopteran-active Bacillus thuringiensis delta-endotoxin polypeptides
and the polynucleotides that encode them, useful for increasing the
insect resistance of plant.
XX
PS Claim 5; Page 139-141; 173pp; English.
XX
CC The sequence represents a B. thuringiensis Lepidopteran-active delta-
endotoxin, crystal protein CryET64. The Lepidopteran-active B.
thuringiensis delta-endotoxin polypeptides may be used as compositions
that are applied to plant crops to protect them from insect damage. The
polynucleotides may be used in the production of transgenic plants that
express the insecticidal polypeptides and consequently have improved
insect resistance compared to non-transformed plants. Monocotyledonous or
dicotyledonous plants may be protected in this way, for example corn,
wheat, soybean, oat, cotton, rice, yre, sorghum, sugarcane, tomato,
tobacco, kapok, flax, potato, barley, turf grass, pasture grass, berry,
fruit, legume, vegetable, ornamental plant, shrub, cactus and/or tree
cell. A wide range of insects (e.g. gypsey moth, looper, tobacco budworm,
cotton leaf perforator and spruce budworm) may be affected by application
of the insecticidal polypeptides (full details given in specification)
XX
SQ Sequence 1228 AA;
```

Query Match 99.7%; Score 6462; DB 4; Length 1228;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1223; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTSNRKNEINEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db :
QY 1 MTSNRKNEINEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db :
QY 61 NIAGILGVLGVPFAGQATASFSPLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120
Db :
QY 61 NIAGILGVLGVPFAGQATASFSPLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120
Db :
QY 121 LARLOGLGDSFRAYQOQSLDLEWLENDLADTRSVLTQYIALELDFLNAMPLFAIRNQVVP 180
Db :
QY 121 LARLOGLGDSFRAYQOQSLDLEWLENDLADTRSVLTQYIALELDFLNAMPLFAIRNQVVP 180
Db :
QY 181 LLMVYAQAANLHLLRLDASLFGSFGSLTSQBIQRYERQVEQTRDYSQVWNTGLN 240
Db :
QY 181 LLMVYAQAANLHLLRLDASLFGSFGSLTSQBIQRYERQVEQTRDYSQVWNTGLN 240
Db :
QY 241 SURGTNAASWVRYNQFRRLDTLGLVLDLVALPFSYDTRTPINTSAQLTREVYTDAGATG 300
Db :
QY 241 SURGTNAASWVRYNQFRRLDTLGLVLDLVALPFSYDTRTPINTSAQLTREVYTDAGATG 300
Db :
QY 301 VNMASSMNNYNAPSFSAIETAVIRSPHLLDLEQLTIFSTSSRSASATRHMTYWRGHTIQ 360
Db :
QY 301 VNMASSMNNYNAPSFSAIETAVIRSPHLLDLEQLTIFSTSSRSASATRHMTYWRGHTIQ 360
Db :
QY 361 SRPIGGLNTSTHGSTNTSINPVRLSPFSRDVYMTESYAGVLLWGIYLEPIHGVPTVRPN 420
Db :
QY 361 SRPIGGLNTSTHGSTNTSINPVRLSPFSRDVYMTESYAGVLLWGIYLEPIHGVPTVRPN 420
Db :
QY 421 FRNPQNTFERGTANTYQSPYESPGLQIKDSETPLETTERPNYESYSHRLSHIGLISQSR 480
Db :
QY 421 FRNPQNTFERGTANTYQSPYESPGLQIKDSETPLETTERPNYESYSHRLSHIGLISQSR 480
Db :
QY 481 VHVPVYVSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVVGPGFTGGDIIRTNVNGS 540
Db :
QY 481 VHVPVYVSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVVGPGFTGGDIIRTNVNGS 540
Db :
QY 541 VLSMGLNFNNTSLQRYRVRVRYAASQTMVLRVTVGGSTTFDQGFPTSMANESLTSQSPR 600
Db :
QY 541 VLSMGLNFNNTSLQRYRVRVRYAASQTMVLRVTVGGSTTFDQGFPTSMANESLTSQSPR 600
Db :
QY 601 FAEFPVGISASGSGTAGISINNAGROTFHFKIEFIPITATFEABYDLERAQEAVALNF 660
Db :
QY 601 FAEFPVGISASGSGTAGISINNAGROTFHFKIEFIPITATFEABYDLERAQEAVALNF 660
Db :
QY 661 TTNTPRLKTDVTDHIDQVSNLVAQLSDFCLDEKRELLKVKYAKRLSDERNLQDPN 720
Db :
QY 661 TTNTPRLKTDVTDHIDQVSNLVAQLSDFCLDEKRELLKVKYAKRLSDERNLQDPN 720
Db :
QY 721 FTSINKQPDFISTNEQSNFTSIHQSEHGWSGENSENITIQEGNDVPKENYVTLPGTFNECY 780
Db :
QY 721 FTSINKQPDFISTNEQSNFTSIHQSEHGWSGENSENITIQEGNDVPKENYVTLPGTFNECY 780
Db :
QY 781 PTYLQKIGESLKYATRYQLRGYIEDSQDLBIYLIRYNAKHETLDVPGTSEWPLSVES 840
Db :
QY 781 PTYLQKIGESLKYATRYQLRGYIEDSQDLBIYLIRYNAKHETLDVPGTSEWPLSVES 840
Db :
QY 841 PIKRGCEPNRCAPHEFWPDLDCRDEKCAHSHHPSLSDIDIGCTDLHENLGVWVFK 900
Db :
QY 841 PIKRGCEPNRCAPHEFWPDLDCRDEKCAHSHHPSLSDIDIGCTDLHENLGVWVFK 900
Db :
QY 901 INTQEGHARLGNLEFIEEKPLLGEALSRYKRAEKKWRDKREKLQLETXRVYTEAKEAADA 960
Db :
QY 901 INTQEGHARLGNLEFIEEKPLLGEALSRYKRAEKKWRDKREKLQLETXRVYTEAKEAADA 960
Db :
QY 961 LFPVDSQYNRLQADTNIGMHAADKLVRIRIAREYLSLSVIFPGVNAIEFEELEGRIITAIS 1020
Db :
QY 961 LFPVDSQYNRLQADTNIGMHAADKLVRIRIAREYLSLSVIFPGVNAIEFEELEGRIITAIS 1020
Db :
QY 1021 LYDARNVKNKGFNNGLACWNKGVHDVQOQSHHRSVLVLPWEAEVSOAVRVCPRGVYL 1080
Db :
QY 1021 LYDARNVKNKGFNNGLACWNKGVHDVQOQSHHRSVLVLPWEAEVSOAVRVCPRGVYL 1080
Db :

QY 1081 RVTAYKEGYGCGVTTIHEIENNTDELKPKNCEEBEVYPTDTGTCNDYTAHQGTAVCNERN 1140
Db :
QY 1081 RVTAYKEGYGCGVTTIHEIENNTDELKPKNCEEBEVYPTDTGTCNDYTAHQGTAVCNERN 1140
Db :
QY 1141 AGYEDAVEVDVTASVNYKPTVEEETDVRDNDHCEYDRGVYVNPPLPAGYMTKELEYPP 1200
Db :
QY 1141 AGYEDAVEVDVTASVNYKPTVEEETDVRDNDHCEYDRGVYVNPPLPAGYMTKELEYPP 1200
Db :
QY 1201 ETDKVMWIEIGETEGKFIVDVSVELLIMEE 1228
Db :
QY 1201 ETDKVMWIEIGETEGKFIVDVSIELLIMEE 1228
Db :
RESULT 3
AAW44321
ID AAW44321 standard; protein; 1227 AA.
XX
AC AAW44321;
XX
DT 27-MAY-1998 (first entry)
XX
DE Bacillus thuringiensis 158C2c toxin.
XX
KW Bacillus thuringiensis; 158C2c; toxin; lepidopteran-active; insect;
KW microbe; transgenic plant; resistant.
XX
OS Bacillus thuringiensis.
XX
FN USS723758-A.
XX
PD 03-MAR-1998.
XX
PF 23-MAY-1995; 95US-00448170.
XX
PR 13-SEP-1991; 91US-00759247.
PR 01-JUN-1993; 93US-00069902.
XX
PA (MYCO) MYCOGEN CORP.
XX
PI Stelman S, Narva KE, Cummings DA, Payne J, Cannon RJ;
XX
DR WPI; 1998-178600/16.
DR N-PSDB; AAV15222.
XX
PT DNA encoding Bacillus thuringiensis toxin proteins - for producing
transgenic plants resistant to attack by lepidopteran pests.
XX
PS Claim 1; Col 33-40; 25pp; English.
XX
CC The present sequence represents a 158C2c toxin from Bacillus
thuringiensis. The present invention describes isolated nucleic acid
sequences encoding toxins active against lepidopteran insects. The
present invention also describes a recombinant host transformed with a
nucleic acid sequence encoding a toxin. The nucleic acid sequence can be
amplified. The product is used for producing transgenic plants resistant
to attack by lepidopteran pests. N.B. The numbers in the sequence listing
do not correspond to the SEQ ID NO.'s mentioned in the specification e.g.
in the specification SEQ ID NO:3, 4, 5 and 6 are said to correspond to
the nucleic acid sequences encoding the toxins, and SEQ ID NO:7, 8, 9 and
10 are said to correspond to the amino acid sequences of the toxins,
whereas in the sequence listing the amino acid sequence is given after
the nucleic acid sequence encoding it (i.e. SEQ ID NO:3 encodes SEQ ID
NO:4)
XX
SQ Sequence 1227 AA;
Query Match 91.5%; Score 5926.5; DB 2; Length 1227;
Best Local Similarity 91.8%; Pred. NO. 0;
Matches 1129; Conservative 36; Mismatches 60; Indels 5; Gaps 3;
QY 1 LTSNRKNEINEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db :
QY 1 LTSNRKNEINEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNIDPFVSVSTVQTGI 60
Db :


```
QY 181 LLVVYAAQANLHLLLRDASLFGSEFGLTSQBIQYRQVQEQTRDYSQVCVWYNTGLN 240
Db 181 LLVVYAAQANLHLLLRDASLFGSEFGLTSQBIQYRQVQEQTRDYSQVCVWYNTGLN 240
QY 241 SLRGNTAAASWVYNQPRDLTLGLDLVALPFSYDTRTPINTSAQLTRVVTDAIGATG 300
Db 241 NLRGNTAAASWVYNQPRDLTLGLDLVALPFSYDTRTPINTSAQLTRVVTDAIGATG 300
QY 301 V--NMASMMWYNNAPSFSAIETAVIRSPHLLDFLEQLTFSTSRWSATRMWYRGHT 358
Db 301 V--NMASMMWYNNAPSFSAIETAVIRSPHLLDFLEQLTFSTSRWSATRMWYRGHT 358
QY 301 APSGFASTWNNAPSFSAIETAVIRSPHLLDFLEQLTFSTSRWSATRMWYRGHT 360
Db 301 APSGFASTWNNAPSFSAIETAVIRSPHLLDFLEQLTFSTSRWSATRMWYRGHT 360
QY 359 IQSPRIQGGINTSTHGSTNTSINPVLSPFSDVWYTESYAGVLWGIYLEPIHGVTVR 418
Db 359 IQSPRIQGGINTSTHGSTNTSINPVLSPFSDVWYTESYAGVLWGIYLEPIHGVTVR 418
QY 361 LESRTIRGSLSTHGNTNTSINPVLQFTSRDVYTESFAGINI--LLTTPVNGVPMAR 418
Db 361 LESRTIRGSLSTHGNTNTSINPVLQFTSRDVYTESFAGINI--LLTTPVNGVPMAR 418
QY 419 ENFRPNQPTFERGTANYSPQYSPGLQXDSLTELPPETTERPNYESYSHRLSHGLISQ 478
Db 419 ENFRPNQPTFERGTANYSPQYSPGLQXDSLTELPPETTERPNYESYSHRLSHGLISQ 478
QY 419 FWRNPLNSL-RGSLLYTIGYTGVTGQLDFSETLPPETTERPNYESYSHRLSNIRLSG 477
Db 419 FWRNPLNSL-RGSLLYTIGYTGVTGQLDFSETLPPETTERPNYESYSHRLSNIRLSG 477
QY 479 SRVHPVSVWTHRSADRTNTISDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVN 538
Db 479 SRVHPVSVWTHRSADRTNTISDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVN 538
QY 478 NTLRAPVSVWTHRSADRTNTISDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVN 537
Db 478 NTLRAPVSVWTHRSADRTNTISDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVN 537
QY 539 GSVLSGLNPNNTSLQRYRVRYAASQTMVLRTVGGSTTFDQGPPTMSANESLTSQS 598
Db 539 GSVLSGLNPNNTSLQRYRVRYAASQTMVLRTVGGSTTFDQGPPTMSANESLTSQS 598
QY 538 GSVLSGLNPNNTSLQRYRVRYAASQTMVLRTVGGSTTFDQGPPTMSANESLTSQS 597
Db 538 GSVLSGLNPNNTSLQRYRVRYAASQTMVLRTVGGSTTFDQGPPTMSANESLTSQS 597
QY 599 FPAEPFVGISAGSGTAGISISNNAGRTTFHDKIEFIPITATFEAEYDLERAQAVNA 658
Db 599 FPAEPFVGISAGSGTAGISISNNAGRTTFHDKIEFIPITATFEAEYDLERAQAVNA 658
QY 598 FPAEPFVGISAGSGTAGISISNNAGRTTFHDKIEFIPITATFEAEYDLERAQAVNA 657
Db 598 FPAEPFVGISAGSGTAGISISNNAGRTTFHDKIEFIPITATFEAEYDLERAQAVNA 657
QY 659 LFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELLEKVKYAKRLSDERNLLQD 718
Db 659 LFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELLEKVKYAKRLSDERNLLQD 718
QY 658 LFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELLEKVKYAKRLSDERNLLQD 717
Db 658 LFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELLEKVKYAKRLSDERNLLQD 717
QY 719 PNFTSINKQDPFISTNEQSNFTSIHQSEHGHWGSENITIQEGNDVFKENYVTLPGTNE 778
Db 719 PNFTSINKQDPFISTNEQSNFTSIHQSEHGHWGSENITIQEGNDVFKENYVTLPGTNE 778
QY 718 PNFTSINKQDPFISTNEQSNFTSIHQSEHGHWGSENITIQEGNDVFKENYVTLPGTNE 777
Db 718 PNFTSINKQDPFISTNEQSNFTSIHQSEHGHWGSENITIQEGNDVFKENYVTLPGTNE 777
QY 779 CYPVLYQKIGSELKAYTRYQLRGVIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
Db 779 CYPVLYQKIGSELKAYTRYQLRGVIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
QY 778 CYPVLYQKIGSELKAYTRYQLRGVIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 837
Db 778 CYPVLYQKIGSELKAYTRYQLRGVIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 837
QY 839 ESPICRCEPNRCAPHFENPDLDCSRDGEKCAHSHHFLSDIDIGCTDLHENLGVWV 898
Db 839 ESPICRCEPNRCAPHFENPDLDCSRDGEKCAHSHHFLSDIDIGCTDLHENLGVWV 898
QY 838 ESPICRCEPNRCAPHFENPDLDCSRDGEKCAHSHHFLSDIDIGCTDLHENLGVWV 897
Db 838 ESPICRCEPNRCAPHFENPDLDCSRDGEKCAHSHHFLSDIDIGCTDLHENLGVWV 897
QY 899 FKIKTOEGHARLGNLEFIEEKEPLLGEALSRVRAEKWRDKREKLQLETKRVYTEAKAV 958
Db 899 FKIKTOEGHARLGNLEFIEEKEPLLGEALSRVRAEKWRDKREKLQLETKRVYTEAKAV 958
QY 959 DALFVDSQYNRLQADNTGMIHAADKLVRHIREAYLSELVTPGVNAIFEELEGRITA 1018
Db 959 DALFVDSQYNRLQADNTGMIHAADKLVRHIREAYLSELVTPGVNAIFEELEGRITA 1018
QY 958 DALFVDSQYNRLQADNTGMIHAADKLVRHIREAYLSELVTPGVNAIFEELEGRITA 1017
Db 958 DALFVDSQYNRLQADNTGMIHAADKLVRHIREAYLSELVTPGVNAIFEELEGRITA 1017
QY 1019 ISLYDARNVKNKGFNNGLACWNVKGVHDVQSHRSVLVPEWAEVSOAVRVCPRGY 1078
Db 1019 ISLYDARNVKNKGFNNGLACWNVKGVHDVQSHRSVLVPEWAEVSOAVRVCPRGY 1078
QY 1018 ISLYDARNVKNKGFNNGLACWNVKGVHDVQSHRSVLVPEWAEVSOAVRVCPRGY 1077
Db 1018 ISLYDARNVKNKGFNNGLACWNVKGVHDVQSHRSVLVPEWAEVSOAVRVCPRGY 1077
QY 1079 ILRVTAKEGYGEGCVTTHIEENNTDELKFNCEEEVYPTDTGTCNDYTAHQGTAVCN 1138
Db 1079 ILRVTAKEGYGEGCVTTHIEENNTDELKFNCEEEVYPTDTGTCNDYTAHQGTAVCN 1138
QY 1078 ILRVTAKEGYGEGCVTTHIEENNTDELKFNCEEEVYPTDTGTCNDYTAHQGTAVCN 1137
Db 1078 ILRVTAKEGYGEGCVTTHIEENNTDELKFNCEEEVYPTDTGTCNDYTAHQGTAVCN 1137
QY 1139 RNAGYEDAYEVDTTASVNYKPYEETVTDVRDNHCEYDRGVNYVPLPAGYMTKELEY 1198
Db 1139 RNAGYEDAYEVDTTASVNYKPYEETVTDVRDNHCEYDRGVNYVPLPAGYMTKELEY 1198
QY 1138 RNAGYEDAYEVDTTASVNYKPYEETVTDVRDNHCEYDRGVNYVPLPAGYMTKELEY 1197
Db 1138 RNAGYEDAYEVDTTASVNYKPYEETVTDVRDNHCEYDRGVNYVPLPAGYMTKELEY 1197
QY 1199 PPETDKVWIEIGETGKFTVDSVELLMEE 1228
Db 1199 PPETDKVWIEIGETGKFTVDSVELLMEE 1227
```

RESULT 5

AAU02046
ID AAU02046 standard; protein; 1227 AA.

XX AAU02046;

XX 29-AUG-2001 (first entry)

XX B. thuringiensis toxic crystal protein, CryET54.

XX Delta endotoxin; Lepidopteran-active; crystal protein; insecticide;
transgenic plant; corn; wheat; soybean; oat; cotton; rice; sorghum;
sugarcane; tomato; tobacco; kapok; flax; potato; barley; turf grass;
pasture grass; berry; fruit; legume; vegetable; ornamental plant; shrub;
cactus; tree cell; gypsy moth; looper; tobacco budworm; spruce budworm;
cotton leaf perforator; CryET54.

XX Bacillus thuringiensis.

OS WO200119859-A2.

XX 22-MAR-2001.

XX 13-SEP-2000; 2000WO-US025361.

XX 15-SEP-1999; 99US-0153995P.

XX (MONS) MONSANTO CO.

XX Baum JA, Chu C, Donovan WP, Gilmer AJ, Rupar MJ;

XX WPI; 2001-281518/29.

XX N-PSDB; AAS02489.

XX Lepidopteran-active Bacillus thuringiensis delta-endotoxin polypeptides
and the polynucleotides that encode them, useful for increasing the
insect resistance of plant.

XX Claim 17; Page 169-173; 173pp; English.

XX The sequence represents B. thuringiensis Lepidopteran-active delta-
endotoxin, crystal protein CryET54. The Lepidopteran-active B.
thuringiensis delta-endotoxin polypeptides may be used as compositions
that are applied to plant crops to protect them from insect damage. The
polynucleotides may be used in the production of transgenic plants that
express the insecticidal polypeptides and consequently have improved
insect resistance compared to non-transformed plants. Monocotyledonous or
dicotyledonous plants may be protected in this way, for example corn,
wheat, soybean, oat, cotton, rice, rye, sorghum, sugarcane, tomato,
tobacco, kapok, flax, potato, barley, turf grass, pasture grass, berry,
fruit, legume, vegetable, ornamental plant, shrub, cactus and/or tree
cell. A wide range of insects (e.g. gypsy moth, looper, tobacco budworm,
cotton leaf perforator and spruce budworm) may be affected by application
of the insecticidal polypeptides (full details given in specification).
Note: The present sequence does not have a cryET number assigned
explicitly in the specification, cryET54 (see table 5, page 56) is not
assigned to any Seq ID number, the index has matched the spare sequence
to the spare cryET number, which may be incorrect

XX Sequence 1227 AA;

Query Match 91.3%; Score 5912.5; DB 4; Length 1227;

Best Local Similarity 91.6%; Pred. No. 0;

Matches 1127; Conservative 36; Mismatches 62; Indels 5; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPARIEDSLCIAEGNNINPLVSASTVQTGI 60

Db 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPARIEDSLCIAEGNNINPLVSASTVQTGI 60

QY 61 NIAGRIILGVLPFAGQIASFYSFLVGLWPRGRDOWEIFLEHVQLINQITENARNTA 120

Db 61 NIAGRIILGVLPFAGQIASFYSFLVGLWPRGRDOWEIFLEHVQLINQITENARNTA 120


```

Db 659 ALFTSTNPRRLKTDVTDYHIDQVSNMVACLSDEFCLDEKRELFKVKYAKRLSDERNLQ 718
Qy 718 DPNETSINKQDPFISTNEQSNFTSIHQSEHGWMGSENITIQEGNDVFKENYVTLPGTFN 777
Db 719 DPNFTFISGQLSFASIDQSNFPIINELSHGWMGSANVTIQEGNDVFKENYVTLPGTFN 778
Qy 778 ECPYLYIKIGESLKAYTRYQLRGVIEDSQLEILYLYRYNAKHETLDPGTESWPLS 837
Db 779 ECPYLYIKIGESLKAYTRYQLRGVIEDSQLEILYLYRYNAKHETLDPGTESWPLS 838
Qy 838 VESPIGRGCPNRCAPHFENWPDSCDRGCKCAHSHHFSIDIDIGCTDLHENLGVWV 897
Db 839 VESPIGRGCPNRCAPHFENWPDSCDRGCKCAHSHHFSIDIDIGCTDLHENLGVWV 898
Qy 898 VFYIKTQEGHARLGNLEFTEKPLGELSRVKRAEKKWRDKRKLQLETKRVYTEAKEA 957
Db 899 VFYIKTQEGHARLGNLEFTEKPLGELSRVKRAEKKWRDKRKLQLETKRVYTEAKEA 958
Qy 958 VDALFVDSQVNRLOADTNIGMIAADKLVHRIREAYLSGELSVPICGVNAEIFELEGRIT 1017
Db 959 VDALFVDSQVNRLOADTNIGMIAADKLVHRIREAYLSGELSVPICGVNAEIFELEGRIT 1018
Qy 1018 AISLYDARNVKNQDFNGLACWNVKGVHDVQOSSHRSVLVPIPEWAEVSQAVRVCPRG 1077
Db 1019 AMSLYDARNVKNQDFNGLACWNVKGVHDVQOSSHRSVLVPIPEWAEVSQAVRVCPRG 1078
Qy 1078 YILRVTAKEGEGCVTHIEENNTDELKFNCEEBEVPTDTGTCNDYTAHQGTAVCN 1137
Db 1079 YILRVTAKEGEGCVTHIEENNTDELKFNCEEBEVPTDTGTCNDYTAHQGTAAACN 1138
Qy 1138 SRNAGVEDAYEDTTASVNYKPTVEEYTTDVRDNCHEYDRGVNYPPLPAGVYTKELE 1197
Db 1139 SRNAGVEDAYEDTTASVNYKPTVEEYTTDVRDNCHEYDRGVNYPPLPAGVYTKELE 1198
Qy 1198 YFPETDKWIEIGETEGKFIVDSEVLLMEE 1228
Db 1199 YFPETDKWIEIGETEGKFIVDSEVLLMEE 1229

```

RESULT 10

```

AAW17699
ID AAW17699 standard; protein; 1229 AA.
AC AAW17699;
XX
XX 25-MAR-2003 (revised)
XX 07-JUL-1997 (first entry)
XX
XX CryET5.
XX
XX CryET5; cryET4; Bacillus thuringiensis; insecticidal crystal protein;
XX ICP; toxin; CryI protein; lepidopteran insect; insecticide.
XX
XX Bacillus thuringiensis.
XX
XX US5616319-A.
XX
XX 01-APR-1997.
XX
XX 30-DEC-1993; 93US-00176865.
XX
XX 29-JUL-1993; 93US-00100709.
XX
XX (MONS ) MONSANTO CO.
XX
XX Gonzalez JM, Donovan WP, Tan Y, Jany CS;
XX
XX WPI; 1997-212077/19.
XX N-PSDB; AAT68434.
XX
XX Bacillus thuringiensis cryET5 gene encoding insecticidal protein - useful
XX for control of lepidopteran pests.
XX

```

XX

Claim 1; Fig 2; 50pp; English.

This sequence represents the cryET5 protein of *Bacillus thuringiensis* (B.t.) isolate EG547. B.t. produces inclusions during sporulation which include insecticidal crystal proteins (ICP). ICP toxins are active in insects only after ingestion. Once ingested, the toxic components disrupt the midgut cells, resulting in cessation of feeding, and eventually death. The CryI proteins produced by B.t. are active against lepidopteran insects. This protein, and the CryET4 protein (see AAW17700) belong to the CryI family of ICPs. The DNA encoding this sequence can be used to transform bacteria, which are useful as insecticides against a wide range of lepidopteran pests, and can be applied to crops, soil and seeds. This protein, or especially its toxic N terminal region, can be expressed in plants, to provide protection against lepidopteran pests. The gene encoding this sequence, or its fragments, can also be used to isolate other similar genes. (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 1229 AA;

Query Match 80.8%; Score 5237.5; DB 2; Length 1229;
 Best Local Similarity 79.9%; Pred. No. 0;
 Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

```

Qy 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Qy 61 NIAGRILGVLPFAGQATASFYSLVGLWPRGRDQWEIFLEHVEQLNQITENARNTA 120
Db 61 NIAGRILGVLPFAGQATASFYSLVGLWPRGRDQWEIFLEHVEQLNQITENARNTA 120
Qy 121 LARLOGLDPSFRAYQOQSLDMLNDRDARTSRVLYTOYIALELDFLAMPPLFAIRNOEVP 180
Db 121 LARLEGLRGVRSYQOALETWLDNRNDARSIIILERYVALELDITTAIFLRIRNEVP 180
Qy 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQEIORYYERQVEQTRDYSVCVEWYNTGLN 240
Db 181 LLMVYQAANLHLLLRDASLFGSEWGMASDVNQYQEQIRYTESSNHCVQWYNTGLN 240
Qy 241 SLRGTTAASWVRYNQFRDLTLGVLDLVALFPSSYDTRTPINTSAQLTREVYTGATG 300
Db 241 NLRGTTAASWVRYNQFRDLTLGVLDLVALFPSSYDTRTPINTSAQLTREVYTGATG 300
Qy 301 V--NMAWNVNNAAPSAIETAVIRSPHLLDPLEQLTIFSTSSRSATRHMTYWRGHT 358
Db 301 APSGFASINWFNNAPSAIETAVIRSPHLLDPLEQLTIFSTSSRSATRHMTYWRGHT 358
Qy 359 IQSRPIGGGLNTSTHGST-NTSINPVRLSPFSDRVYVWTESYAGVLLWGIYLEPHGVPTV 417
Db 361 LNFRPIGGGLNTSTHGST-NTSINPVRLSPFSDRVYVWTESYAGVLLWGIYLEPHGVPTV 417
Qy 418 RFNFRNQNTFERGTANYQPYSPGLQKDSSTELPPETTERPNYESYSHRLSHLGLIS 477
Db 419 RFNFINQNIYERGATTYSQYQGVGILQFDSSTELPPETTERPNYESYSHRLSHLGLII 478
Qy 478 QSRVHVPEVYSWTHRSADRTNTISSDSITQIPVKSFNLSGTSVSGPGFTGDDIIRTNV 537
Db 479 GNTLRAPVYSWTHRSADRTNTIGPNRITQIPVKAHLNLSHGVTVVGGPGFTGDDIIRTN 538
Qy 538 NGSVLSGLNFNNTSIQRYVRVRYAASQTMVLRLVTVGGSSTTFDQGFPSMTSANEISLTSQ 597
Db 539 TGTFGDIRLINVPLSQRYVRVRYASTTDLQFFTRINGTIVNIGNFSRTMRGDNLEVR 598
Qy 598 SFRFAEFPFGISAGSQTAGISINNAGQTHFDKIEFIPITATPEAYDYLERAQEAVN 657
Db 599 SFRTAGFSTFPFLNAQSTFTLGAOSFSQEQYVIDRVEFVPAEVTFEAYDYLERAKAVN 658
Qy 658 ALPTNTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELFKVKYAKRLSDERNLQ 717
Db 659 ALPTNTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELFKVKYAKRLSDERNLQ 718
Qy 718 DPNFTSINKQDPFISTNEQSNFTSIHQSEHGWMGSENITIQEGNDVFKENYVTLPGTFN 777

```

719	Db	DPNFTFISGQSPASIDQGQSNFFPSINSELSBHGWMGSANVTIQEGNDVFKENYYTLPGTFN	778
778	Qy	ECYPTYLYQKIGESSELKAYTRYQLRGYIEDSQDLEIYLIRYNKAGHETLDVPGTESVWPLS	837
779	Db	ECYPNYLYQKIGESSELKAYTRYQLRGYIEDSQDLEIYLIRYNKAGHETLDVPGTDSLWPLS	838
838	Qy	VESPIGRGCBPNRCAPHFENWPNLDCCSRGDEKCAHHSHHFSLDIDTGCITDLHENTLGVVY	897
839	Db	VESPIGRGCBPNRCAPHFENWPNLDCCSRGDERCAHHSHHFTLIDVGCITDLHENTLGVVY	898
898	Qy	VFKIKTOEGHARLGNLFSEIPEKPLIGALSRLVKRAEKKWRDKKEKLOLETKRYVYTRAKEA	957
899	Db	VFKIKTOEGHARLGNLFSEIPEKPLIGALSRLVKRAEKKWRDKKEKLOLETKRYVYTRAKEA	958
958	Qy	VDALFVDSQYNRLQADTNIGMIIHAADKLVHRIREAYLSLSVPIGVNABIFEELGRIT	1017
959	Db	VDALFVDSQYDQLQADTNIGMIIHAADKLVHRIREAYLSLVPVPGVNAIFEELGHIIT	1018
1018	Qy	AISLYDARNVYKQDFFNGLACWNVKGHVDVQOSHRSVLVPIPEWAEVVSQAVRVCPGRG	1077
1019	Db	AMSLYDARNVYKQDFFNGLTCWNVKGHVDVQOSHRSDLVPIPEWAEVVSQAVRVCPGRG	1078
1078	Qy	YILRVVTAYPEGKGEGCVTHIEINNTDELKPKNCEBEEVYPTDTGTCNDYTAHQGTAVCN	1137
1079	Db	YILRVVTAYPEGKGEGCVTHIEINNTDELKPKNCEBEEVYPTDTGTCNDYTAHQGTAAACN	1138
1138	Qy	SRNAGYEDAYEVDTTASVNVKPTYEEETYYTDVRRDNHCEYDRGVVNYPPLPAGVMTKELE	1197
1139	Db	SRNAGYEDAYEVDTTASVNVKPTYEEETYYTDVRRDNHCEYDRGVVNYPPVPAGVYKELE	1198
1198	Qy	YFPETDKVWIEIGETGEKFIVDSEVLLLMEE	1228
1199	Db	YFPETDKVWIEIGETGEKFIVDSEVLLLMEE	1229

RESULT 11

AAW87633	AAW87633 standard; protein; 1229 AA.
XX	
XX	
XX	
XX	
XX	
DT	03-MAR-1999 (first entry)
XX	
DE	CryET5 protein sequence.
XX	
XX	CryET4; Bacillus thuringiensis strain EG5847; crystal toxin; CryET5;
KW	Lepidoptera; Helicoverpa zea; resistant strain; Plutella xylostella;
KW	Spodoptera exigua; S. frugiperda; S. frugiperda; Trichoplusia ni.
XX	
XX	Bacillus thuringiensis.
OS	
XX	
XX	US5854053-A.
PN	
XX	
PD	29-DEC-1998.
XX	
XX	
PF	06-JAN-1997; 97US-00779046.
XX	
XX	
XX	29-JUL-1993; 93US-00100709.
PR	30-DEC-1993; 93US-00176865.
PR	07-JUN-1995; 95US-00474038.
XX	
XX	
PA	(ECOG-) ECOGEN INC.
XX	
PI	Donovan WP, Gonzalez JM;
XX	
XX	WPI; 1999-094915/08.
DR	N-PSDB; AAV83927.
DR	
XX	
XX	
PT	New strains EG5847 and EG10368 of Bacillus thuringiensis - producing
PT	crystal toxins active against Lepidoptera, particularly Helicoverpa zea,
PT	Plutella xylostella and spodoptera species.
XX	
XX	

QY 838 VESPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLDIDIGCTDLHENLGVWV 897
 Db 839 VESPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLDIDIGCTDLHENLGVWV 898
 QY 898 VFQIKTQEGHARLGNLEFIEEKPLLEALSRVKRAEKKWRDKREKQLQLETKRVVTEAKEA 957
 Db 899 VFQIKTQEGHARLGNLEFIEEKPLLEALSRVKRAEKKWRDKREKQLQLETKRVVTEAKEA 958
 QY 958 VDPAFVDSQYNLOADNTNIGMTHAADKLVHRIEAYLSGELSVPICVNAEIEFEELGRIIT 1017
 Db 959 VDPAFVDSQYNLOADNTNIGMTHAADKLVHRIEAYLSGELSVPICVNAEIEFEELGRIIT 1018
 QY 1018 AISLYDARNVKNQGFNNGLACWNVKGVQVQSHRSVLVTPWEAEVSQAVRVCPRG 1077
 Db 1019 ANSLYDARNVKNQGFNNGLTCWNVKGVQVQSHRSVLVTPWEAEVSQAVRVCPRG 1078
 QY 1078 YILRVTAKEGEGECVTIHEIENNTDELKFNCEEEVYPTDGTCTNDYTAHQGTACN 1137
 Db 1079 YILRVTAKEGEGECVTIHEIENNTDELKFNCEEEVYPTDGTCTNDYTAHQGTACN 1138
 QY 1138 SRNAGVEDAYEVDTTASVNYKPTVEEETVTVRRDNHCEYDGVNPPPLPAGVWTKELE 1197
 Db 1139 SRNAGVEDAYEVDTTASVNYKPTVEEETVTVRRDNHCEYDGVNPPVPGVYVTKELE 1198
 QY 1198 YPPTDKVWIEIGETEGKPIVDSVELLMEE 1228
 Db 1199 YPPTDKVWIEIGETEGKPIVDSVELLMEE 1229

RESULT 12

AAV30923

ID AAV30923 standard; protein; 1229 AA.

XX AC AAV30923;

XX DT 18-OCT-1999 (first entry)

XX DE B. thuringiensis cryET5 partial protein.

XX KW Toxin; cryET4; cryET5; insecticidal; Lepidoptera; transformed plant;

XX OS Bacillus thuringiensis.

XX FH Key

XX FT Protein

XX FT 1. 1229

XX FT /notes="partial protein sequence. No ATG start codon

XX FT given"

XX PN US5942658-A.

XX PD 24-AUG-1999.

XX PP 24-JUN-1997; 97US-00881340.

XX PR 29-JUL-1993; 93US-00100709.

XX PR 30-DEC-1993; 93US-00176865.

XX PR 07-JUN-1995; 95US-00474038.

XX PA (MONS) MONSANTO CO.

XX XX Gonzalez JM, Jany CS, Tan Y, Donovan WP;

XX XX WPI; 1999-493544/41.

XX XX N-PSDB; AAZ09160.

XX XX Transformed plant comprising insecticidal crystal proteins.

XX XX Claim 3; Fig 2A-J; 50pp; English.

XX XX This invention describes novel transformed plants containing Bacillus

XX CC thuringiensis strain EG5847 insecticidal crystal protein genes cryET4 or

XX CC cryET5. cryET4 and cryET5 are novel toxin genes which produce

CC insecticidal proteins with activity against a broad spectrum of insects
 CC of the order Lepidoptera. This sequence represents the cryET5 protein
 CC described in the method of the invention

XX SQ Sequence 1229 AA;

Query Match 80.8%; Score 5237.5; DB 2; Length 1229;
 Best Local Similarity 79.9%; Pred. No. 0;
 Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVSNHSTQWLDSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60

Db 1 LTSNRKNEIINALSIPAVSNHSTQWLDSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60

QY 61 NIAGRILGVLPFAGQIASFYFLGELWPGRDOWEIFLEHVEQLINQQTENARNTA 120

Db 61 NIAGRILGVLPFAGQIASFYFLGELWPGRDOWEIFLEHVEQLINQQTENARNTA 120

QY 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSVLYTQYIALELDLFNAMPLFAIRNOEVP 180

Db 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSVLYTQYIALELDLFNAMPLFAIRNOEVP 180

QY 181 LLMVYAAQANLHLLLDASLFGSFGELTQBIQRYERQVOTRDYSYCVENYNTGLN 240

Db 181 LLMVYAAQANLHLLLDASLFGSFGELTQBIQRYERQVOTRDYSYCVENYNTGLN 240

QY 241 SLRGTNAAWRYNQFRRLDTLGLVDLVALPSPYDTRTPYNTSAQLTREYVTDGATG 300

Db 241 SLRGTNAAWRYNQFRRLDTLGLVDLVALPSPYDTRTPYNTSAQLTREYVTDGATG 300

QY 301 V--NNASWNNYNNAPSAIETAVIRSPHLLDFLEQLTIFSTSSRSWATRHMTYWRGT 358

Db 301 V--NNASWNNYNNAPSAIETAVIRSPHLLDFLEQLTIFSTSSRSWATRHMTYWRGT 358

QY 359 IQSRPIGGGLNTSTHGST--NTSINPVLSPFSRDVYVWTSYAGVLLWYIPIPIGVPTV 417

Db 359 IQSRPIGGGLNTSTHGST--NTSINPVLSPFSRDVYVWTSYAGVLLWYIPIPIGVPTV 417

QY 418 RNFNPNQNTFERGTANYSPQVSGQLKDSBELPETTERPNYESYSHRLSHIGLIS 477

Db 418 RNFNPNQNTFERGTANYSPQVSGQLKDSBELPETTERPNYESYSHRLSHIGLIS 477

QY 478 QSRVHPVYVWTHRSADRTNTISSDITQIPLVSKFNLSGTSVVSQVSGFGTGGDIIRNV 537

Db 478 QSRVHPVYVWTHRSADRTNTISSDITQIPLVSKFNLSGTSVVSQVSGFGTGGDIIRNV 537

QY 538 NGSVLSMGLNFNTSLQRYVRVRYAASQTMVLRVTVGGSTTFDQGFSTMSANESLTSQ 597

Db 538 NGSVLSMGLNFNTSLQRYVRVRYAASQTMVLRVTVGGSTTFDQGFSTMSANESLTSQ 597

QY 598 SFRFAEPVGISASGQTAGISINNAGRTQFHFKIEFIPITATFPAEYDLERAQEA 657

Db 598 SFRFAEPVGISASGQTAGISINNAGRTQFHFKIEFIPITATFPAEYDLERAQEA 657

QY 658 ALFTNTNPRRLKTDVTDYHIDOVSNLVACLSDEFCLDEKRELLKVKYAKELSDERMLQ 717

Db 658 ALFTNTNPRRLKTDVTDYHIDOVSNLVACLSDEFCLDEKRELLKVKYAKELSDERMLQ 717

QY 718 DPNFTSINKQDPFISTNEQNFSTSIHQSEHGMWGSNITIQEGNDVFKENYVTLPGTFN 777

Db 718 DPNFTSINKQDPFISTNEQNFSTSIHQSEHGMWGSNITIQEGNDVFKENYVTLPGTFN 777

QY 778 ECYPTLYQKIGSELKAYTRYQLRGYIEDSDOLEIYLIRYNAKHETLDVPGTESVWPLS 837

Db 778 ECYPTLYQKIGSELKAYTRYQLRGYIEDSDOLEIYLIRYNAKHETLDVPGTESVWPLS 837

QY 838 VESPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLDIDIGCTDLHENLGVWV 897

Db 838 VESPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLDIDIGCTDLHENLGVWV 897

QY 898 VFQIKTQEGHARLGNLEFIEEKPLLEALSRVKRAEKKWRDKREKQLQLETKRVVTEAKEA 957

Db 898 VFQIKTQEGHARLGNLEFIEEKPLLEALSRVKRAEKKWRDKREKQLQLETKRVVTEAKEA 957

QY 958 VDALFVDSQYNRLQADTNIGMIHAADKLVHRIEAYLSLSVIPGVNAIPEELGRIIT 1017
 DB 959 VDALFVDSQYDQDLQADTNIGMIHAADKLVHRIEAYLSLSELPVPGVNAIPEELGHIIT 1018
 QY 1018 AISLYDARNVKNNGDPNGLACWNVKGVHVDVQOSHRSVLVITPEWEAEVSQAVRVCPRG 1077
 DB 1019 AMSLYDARNVKNNGDPNGLTCWNVKGVHVDVQOSHRSDLVITPEWEAEVSQAVRVCPRG 1078
 QY 1078 YILRTVAYKEGEGCVTTHETENNTDELKPKNCBEEVYPTDGTCDNDYTAHQGTAVCN 1137
 DB 1079 YILRTVAYKEGEGCVTTHETENNTDELKPKNCBEEVYPTDGTCDNDYTAHQGTACN 1138
 QY 1138 SRNAGYEDAYEDVTTASVNYKPYEETVTVRRDNHCEYDRGVYNYPLPAGYMTKELE 1197
 DB 1139 SRNAGYEDAYEDVTTASVNYKPYEETVTVRRDNHCEYDRGVYNYPPVAGYVTKELE 1198
 QY 1198 YPPTDKWVIEIGETEGKFIIVDSVELLMEE 1228
 DB 1199 YPPTDTVWIEIGETEGKFIIVDSVELLMEE 1229

RESULT 13

ADK98479

ID ADK98479 standard; protein; 1229 AA.

AC ADK98479;

XX 03-JUN-2004 (first entry)

DT

XX B thuringiensis cryIbB partial protein sequence SeqID2.

DE

XX insecticidal protein; plant; pesticide; gene therapy;

KW

XX lepidopteran insect pest; transgenic plant;

KW

XX insect infestation resistance; monocot; dicot; cryIbB.

XX

OS Bacillus thuringiensis.

XX

XX WO2004020636-A1.

PN

XX 11-MAR-2004.

PD

XX 26-AUG-2003; 2003WO-US026510.

PF

XX 29-AUG-2002; 2002US-0407428P.

PR

XX (MONS) MONSANTO TECHNOLOGY LLC.

PA

XX (BOGD/) BOGDANOVA N N.

PA

XX (ROMA/) ROMANO C P.

XX

XX Bogdanova NN, Romano CP;

PI

XX WPI; 2004-269221/25.

XX

XX DR N-P5DB; ADK98478.

DR

New polynucleotide sequence optimized for expression of an insecticidal protein in a plant, useful in the control of Lepidoptera insect pests, and for producing transgenic plants with the ability to resist insect infestations.

PT

XX Claim 7; SEQ ID NO 2; 138pp; English.

PS

XX This invention relates to a novel polynucleotide sequence optimised for expression of an insecticidal protein in a plant. The invention may be useful for the production of pesticides whilst the disclosed sequences may be used for gene therapy. The polynucleotide sequence and methods are useful in the control of lepidopteran insect pests, and for producing transgenic plants with the ability to resist insect infestations. The invention provides polynucleotide sequences with enhanced, improved and optimised expression in monocot and dicot plant species. The present sequence is that of the (partial) B thuringiensis cryIbB protein which is related to the invention.

CC

XX

SQ Sequence 1229 AA;

Query Match 80.8%; Score 5237.5; DB 8; Length 1229;

Best Local Similarity 79.9%; Pred. No. 0;

Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSDPARIEDSLCIAEGNNINPLVSASTVQTGI 60
 DB 1 LTSNRKNEIINALSIPAVSNHSTQMDLSDPARIEDSLCIAEGNNINPLVSASTVQTGI 60
 QY 61 NIAGRIILGVLPAGQIASFSLVGLMWRGRDQWEIFLEHVQLNQITENARNTA 120
 DB 61 NIAGRIILGVLPAGQIASFSLVGLMWRGRDQWEIFLEHVQLNQITENARNTA 120
 QY 121 LARLQGLGDSFRAVQOSLEDMLENRDDARTSVLTVYIALELDPFLNAPLPAIRNOEVP 180
 DB 121 LARLQGLGDSFRAVQOSLEDMLENRDDARTSVLTVYIALELDPFLNAPLPAIRNOEVP 180
 QY 181 LLWVYAQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSDYCVWEVNTGLN 240
 DB 181 LLWVYAQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSDYCVWEVNTGLN 240
 QY 241 SLRGTTAAASVRYNQFRRLDLGLVLDLVALPSPYDTRTYPIINTSAQLTREYVTDAGATG 300
 DB 241 NLRGTNAESWLRVYNQFRRLDLGLVLDLVALPSPYDTRTYPIINTSAQLTREYVTDIGRTN 300
 QY 301 V--NWASMNWYNNNAPSFAIETAVIRSPHLLDLEQLTIFSTSSRWASATRMVWGRHT 358
 DB 301 APSGFASNTWNNNAPSFAIETAVIRSPHLLDLEQLTIFSTSSRWASATRMVWGRHT 360
 QY 359 IQSRPIGGGLNTSTHGST--NTSINPVRLSPPSRDVTYVWTSYAGVLLWGLYLEPIHGVPV 417
 DB 361 LNFRIPIGGGLNTSTHGST--NTSINPVRLSPPSRDVTYVWTSYAGVLLWGLYLEPIHGVPV 418
 QY 418 RFNFRNPONTFERTGANYQSPYSPGLQKQDSELTPELTPETTERPNYRYSRHLSHIGLIS 477
 DB 419 RFNFINQNIYERGATTSYQPYQGVIGLQKQDSELTPELTPETTERPNYRYSRHLSHIGLII 478
 QY 478 QSRVHVVPVYSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVSVPFGTGGDIIRTNV 537
 DB 479 GNTLRAPVYSWTHRSADRTNTIGPNRITQIPLVKNLHSGTVVVGSGFTGGDILRTN 538
 QY 538 NGSVLSMGLNFNTSLQRYRVRYAASQTMVLRTVTVGGSTTFDQGPSTMSANESLTSQ 597
 DB 539 TGTFGDIRLNLNINPLSQRYRVRYASTDITQFFTRINGTIVNIGFSTRWNRGDNLEVR 598
 QY 598 SFRPAEFPVGISASGQTAGISISNNAGQTFHFQKIEFIPITATFEAYDLEAQAQAVN 657
 DB 599 SFRTAGFSTFPNPLNAQSTFTLGAQSFNSQEVYIDRVFVPAEVTFEAYDLEAQAQAVN 658
 QY 658 ALFTNTNPRRLKTDVTDYHIDVSNLVACLSDEECLDEKRELLKVKYAKLSDERNLLQ 717
 DB 659 ALFTNTNPRRLKTDVTDYHIDVSNLVACLSDEECLDEKRELLKVKYAKLSDERNLLQ 718
 QY 718 DPNTFTSINKQDPDFTSTNEQSNFTSIHQSEHGWGMSNITIQEGNDVFKENYVTLPGTFN 777
 DB 719 DPNTFTSINKQDPDFTSTNEQSNFTSIHQSEHGWGMSNITIQEGNDVFKENYVTLPGTFN 778
 QY 778 ECPYTYLYQKIGSELSKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTSVWPLS 837
 DB 779 ECPYTYLYQKIGSELSKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTSVWPLS 838
 QY 838 VESPIGRGPNRCAPHFENWPDLDCCRCRGEKCAHSHHPSLDIDGCTDLHENLGVWV 897
 DB 839 VESPIGRGPNRCAPHFENWPDLDCCRCRGEKCAHSHHPSLDIDGCTDLHENLGVWV 898
 QY 898 VFKIKTOEGHARLGNLFIEBKPLLGALSRVKAEEKWRDKREKQLQETKRVYTEAKEA 957
 DB 899 VFKIKTOEGHARLGNLFIEBKPLLGALSRVKAEEKWRDKREKQLQETKRVYTEAKEA 958
 QY 958 VDALFVDSQYNRLQADTNIGMIHAADKLVHRIEAYLSLSVIPGVNAIPEELGRIIT 1017
 DB 959 VDALFVDSQYDQDLQADTNIGMIHAADKLVHRIEAYLSLSELPVPGVNAIPEELGHIIT 1018

Db 1021 MSLEYDARNVXKNGDFNNGJTCWNKGVHDVQOQSHHRSDLVPEWEAEVSAQVRVCPGRGY 1080
 QY 1079 ILRVAYKEGYGEGCVTHIEINNTDELKFKNCBEEVYPTDTGTCNDYTAHQGTAVCNS 1138
 Db 1081 ILRVAYKEGYGEGCVTHIEINNTDELKFKNCBEEVYPTDTGTCNDYTAHQGTAAACS 1140
 QY 1139 RNAGVEDAYEVDTTASVNYKPYEBETVYDVRDNHCEYDRGVVNYPLPAGYMTKELEY 1198
 Db 1141 RNAGVEDAYEVDTTASVNYKPYEBETVYDVRDNHCEYDRGVVNYPPVAGYVTKLEY 1200
 QY 1199 PPTDKWTEIGETEGKFIVDSEVILLMBE 1228
 Db 1201 PPTDTVWIEIGETEGKFIVDSEVILLMBE 1230

RESULT 15

ADK98489
 ID ADK98489 standard; protein; 1230 AA.

AC ADK98489;

DT 03-JUN-2004 (first entry)

DE B thuringiensis cryIbB-related expression cassette protein SeqID12.

XX insecticidal protein; plant; pesticide; gene therapy;
 KW lepidopteran insect pest; transgenic plant;
 KW insect infestation resistance; monocot; dicot; cryIbB.

XX Bacillus thuringiensis.
 OS Synthetic.

XX WO2004020636-A1.

XX 11-MAR-2004.

XX 26-AUG-2003; 2003WO-US026510.

XX 29-AUG-2002; 2002US-0407428P.

XX (MONS) MONSANTO TECHNOLOGY LLC.
 PA (BOGD/) BOGDANOVA N N.
 PA (ROMA/) ROMANO C P.

XX Bogdanova NN, Romano CP;

XX WPI: 2004-269221/25.

DR N-P5DB; ADK98488.

XX New polynucleotide sequence optimized for expression of an insecticidal
 PT protein in a plant, useful in the control of Lepidoptera insect pests,
 PT and for producing transgenic plants with the ability to resist insect
 PT infestations.

XX Claim 7; SEQ ID NO 12; 138pp; English.

PS This invention relates to a novel polynucleotide sequence optimised for
 CC expression of an insecticidal protein in a plant. The invention may be
 CC useful for the production of pesticides whilst the disclosed sequences
 CC may be used for gene therapy. The polynucleotide sequence and methods are
 CC useful in the control of lepidopteran insect pests, and for producing
 CC transgenic plants with the ability to resist insect infestations. The
 CC invention provides polynucleotide sequences with enhanced, improved and
 CC optimised expression in monocot and dicot plant species. The present
 CC sequence is that of a B thuringiensis cryIbB expression cassette protein
 CC which is related to the invention.

XX Sequence 1230 AA;

Query Match 80.8%; Score 5233.5; DB 8; Length 1230;
 Best Local Similarity 79.8%; Pred. No. 0;
 Matches 982; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

QY 2 TSNRKNENEIINALSIPAVSNHSTOMDLSDPARIEDSLCAEGNINNPLVSASTVQTGIN 61
 Db 3 TSNRKNENEIINALSIPVSNPSTOMNLSDPARIEDSLCAEVNNDIPVSASTVQTGIN 62
 QY 62 IAGRTILGVLGVPPAGQIASFYSFLVGLWPGRDQWEIPLFHVQLINQOITENARNTAL 121
 Db 63 IAGRTILGVLGVPPAGQIASFYSFLVGLWPGRDQWEIPLFHVQLINQOITENARNTAI 122
 QY 122 ARLOGLGDSFRAYCOOSLEDWLENRDADTRSVLYTQYIALELDPLANPLFAIRNOEVL 181
 Db 123 ARLEGLGRGYSYQOALETDWLDNRNARSIIILERYVALELDITTAIPLFRIRNEVPL 182
 QY 182 LMVYAAQANLHLLLRDASLFGSBSFGLTSQIRYRQYRQVEQTRDYSYCVWEYNTGLNS 241
 Db 183 LMVYAAQANLHLLLRDASLFGSEWMASSDVNQYQEQIRYTBYSNHCYQWYNTGLNN 242
 QY 242 LRGTNAASVRYNOFRDLITGLVDLVALFYSYDTRTYPIINTSAQLTREVYDAIGATGV 301
 Db 243 LRGTNAESWLYNOFRDLITGLVDLVALFYSYDTRTYPIINTSAQLTREVYDPIGRNA 302
 QY 302 --NMASNNVNNNAPSFAIETAVIRSPHLADPLEOLTIPTSSRWSATRWYWRGHTI 359
 Db 303 PSGFASTWNNNAPSFAIEAIFRPPHLLDFPEQLTIYSASSRWSSTOHNNVWGHRL 362
 QY 360 QSRPIGGGLNTSTHGST-NTSINPVRLSFFSRDVVWTSYAGVLIWGIYLPBPIHGVPVTR 418
 Db 363 NFRPIGGTLNTSTOGLTNTSINPVLTQFTSRDVYRTESNAGTNI--LFTTPVNGVPPAR 420
 QY 419 NFRNPQNTFERGTANYSPQYESPOLQKDSLETPPETTERPNYESYSHRLSHGLISQ 478
 Db 421 NFNINPQNIYERGATYSQPYQGVGIQIFDSETELPETTERPNYESYSHRLSHGLIIG 480
 QY 479 SRVHVYVSWTHRSADRTNTSSDSITQIPLVKSFLNLSGTSVVSGRPTCGDIIIRTVN 538
 Db 481 NTLRAPVYVSWTHRSADRTNTIGPNRITQIPLVKALNLHSGVTYVGGPGFTGDIILRRTNT 540
 QY 539 GSVLSMGLNNTSLQRYRVRVRAAASQTMVLVTVGGSTTFDQGFPPSTMA NESLTSQS 598
 Db 541 GTFGDIRLINVPLSQRVVRIRYASTDLOFFTRINGTTVNI GNFSYTMRGDNLEYS 600
 QY 599 FRFAEPVGISASGQTAGISISNNAGRQTFPHKIBFIPITATPEABYDLERAQAVNA 658
 Db 601 FRTAGFSTPFNLAQSTFTLGAQSFNSQEVYIDRVEFVPAEVTPEABYDLERAQAVNA 660
 QY 659 LFTNTPRRLKTDVTDHIDQVSNLVACLSEFCLEKRELVKVKAKRLSDRENLLQD 718
 Db 661 LFTSTNPRRLKTDVTDHIDQVSNVACLSEFCLEKRELVKVKAKRLSDRENLLQD 720
 QY 719 PNFTSINKQPDFISTNEQSNTSIHQSEHGWGSGENITIQEGNDVFKENTVTLPGTNE 778
 Db 721 PNFTFISQLSFASIDQSNFPPSINELSEHGWGSGANVTIQEGNDVFKENTVTLPGTNE 780
 QY 779 CYPYLYQKIGESBELKAYTRYQLRGYIEDSDLEIYLIRYNKHEHTLDVPGTESVWPLSV 838
 Db 781 CYPNLYQKIGESBELKAYTRYQLRGYIEDSDLEIYLIRYNKHEHTLDVPGTDSLWPLSV 840
 QY 839 ESPIGRCENRCAHPHEWNPDLDCSDGCKCAHSHHPSLDIDIGCTDLHENLGVWV 898
 Db 841 ESPIGRCENRCAHPHEWNPDLDCSDGCKCAHSHHPSLDIDIVGCTDLHENLGVWV 900
 QY 899 FKIKTQSGHARLGNLEFTEEKPLIGEALSRVKRAEKWRDKREKLQLETKRVYTEAKAV 958
 Db 901 FKIKTQSGHARLGNLEFTEEKPLIGEALSRVKRAEKWRDKREKLQLETKRVYTEAKAV 960
 QY 959 DALFVDSQYNRLQADTNIGMIHAADKLVRHIREAYLSLSVPIGVNABIFEELEGRIITA 1018
 Db 961 DALFVDSQYDQDQADTNIGMIHAADKLVRHIREAYLSLSVPIGVNABIFEELEGRIITA 1020
 QY 1019 ISLYDARNVXKNGDFNNGJTCWNKGVHDVQOQSHHRSDLVPEWEAEVSAQVRVCPGRGY 1078
 Db 1021 MSLEYDARNVXKNGDFNNGJTCWNKGVHDVQOQSHHRSDLVPEWEAEVSAQVRVCPGRGY 1080
 QY 1079 ILRVAYKEGYGEGCVTHIEINNTDELKFKNCBEEVYPTDTGTCNDYTAHQGTAVCNS 1138

Db	1081	ILRVAYKEGYGECVTIHEIENNTDELKPKNCEEEVYPTDTGTCNDYTAHQGTAAACNS	1140
Qy	1139	RNAGYEDAYEVDTTASVNYKPTYEEETDVRDRNHCEYDRGVNYPPLPAGYMTKELEY	1198
Db	1141	RNAGYEDAYEVDTTASVNYKPTYEEETDVRDRNHCEYDRGVNYPPLPAGYMTKELEY	1200
Qy	1199	FPETDKWIEIGETEGKFIVDSVELLMEE	1228
Db	1201	FPETDKWIEIGETEGKFIVDSVELLMEE	1230

Search completed: November 21, 2004, 17:00:30
Job time : 128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 16:20:51 ; Search time 43 Seconds
(without alignments)
1893.919 Million cell updates/sec

Title: US-10-614-524-2
Perfect score: 6479
Sequence: 1 LTSNRKNEIINALSIPAV.....IGETGKFVDSVEILLMEER 1228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6462	99.7	1228	4	US-09-661-322A-38
2	5926.5	91.5	1227	1	US-08-448-170-8
3	5926.5	91.5	1227	3	US-08-961-803-9
4	5912.5	91.3	1227	4	US-09-661-322A-63
5	5659.5	87.4	1207	1	US-07-951-715A-7
6	5659.5	87.4	1207	2	US-08-459-448A-7
7	5659.5	87.4	1207	3	US-08-459-595A-7
8	5659.5	87.4	1207	3	US-08-459-504B-7
9	5659.5	87.4	1207	3	US-08-459-444-7
10	5659.5	87.4	1207	3	US-09-053-549-8
11	5659.5	87.4	1207	3	US-09-547-422-7
12	5659.5	87.4	1207	4	US-09-988-462-7
13	5436.5	83.9	1227	3	US-09-053-549-2
14	5237.5	80.8	1229	1	US-08-100-709-4
15	5237.5	80.8	1229	1	US-08-176-865-4
16	5237.5	80.8	1229	1	US-08-474-038-4
17	5237.5	80.8	1229	2	US-08-779-046-4
18	5237.5	80.8	1229	2	US-08-881-340-4
19	5108	78.8	1186	3	US-09-178-252-23
20	5108	78.8	1186	4	US-09-826-660-23
21	3977.5	61.4	1174	1	US-08-040-751-3
22	3977.5	61.4	1174	1	US-08-291-368-2
23	3977.5	61.4	1174	2	US-08-962-190-2
24	3977.5	61.4	1174	5	PCT-US95-10310-2
25	3977.5	61.4	1174	6	5164180-4
26	3791	58.5	1176	1	US-08-434-823-2
27	3791	58.5	1176	1	US-08-457-366-2

28	3579	55.2	1170	1	US-08-032-364-2	Sequence 2, Appli
29	3556	54.9	1167	1	US-08-100-709-2	Sequence 2, Appli
30	3556	54.9	1167	1	US-08-176-865-2	Sequence 2, Appli
31	3556	54.9	1167	1	US-08-474-038-2	Sequence 2, Appli
32	3556	54.9	1167	2	US-08-779-046-2	Sequence 2, Appli
33	3556	54.9	1167	2	US-08-881-340-2	Sequence 2, Appli
34	3502.5	54.1	1189	3	US-09-314-093-59	Sequence 59, Appl
35	3502.5	54.1	1189	3	US-09-314-093-59	Sequence 59, Appl
36	3502.5	54.1	1189	4	US-09-337-635-59	Sequence 59, Appl
37	3502.5	54.1	1189	4	US-09-337-280-59	Sequence 59, Appl
38	3500.5	54.0	1189	2	US-08-980-071-2	Sequence 2, Appli
39	3500.5	54.0	1189	2	US-08-757-536-2	Sequence 2, Appli
40	3500.5	54.0	1189	3	US-09-314-093-2	Sequence 2, Appli
41	3500.5	54.0	1189	3	US-09-250-848-2	Sequence 2, Appli
42	3500.5	54.0	1189	3	US-09-251-885-2	Sequence 2, Appli
43	3500.5	54.0	1189	3	US-09-337-635-2	Sequence 2, Appli
44	3500.5	54.0	1189	4	US-09-337-280-2	Sequence 2, Appli
45	3496.5	54.0	1189	1	US-08-602-737-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-661-322A-38
; Sequence 38, Application US/09661322A
; Patent No. 6593293
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rel
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Ruper, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Compos
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: MECO201
; CURRENT APPLICATION NUMBER: US/09/661.322A
; CURRENT FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-09-661-322A-38

Query Match	99.7%	Score 6462;	DB 4;	Length 1228;
Best Local Similarity	99.6%	Pred. No. 0;		
Matches 1223;	Conservative	3;	Mismatches	2; Indels 0; Gaps 0;
Qy	1	LTSNRKNEIINALSIPAVSNHSTQMDLSPDARTEDSLCIAEGNNINPLYSASTVQTGI	60	
Db	1	MTSRKNEIINALSIPAVSNHSTQMDLSPDARTEDSLCIAEGNNINPLYSASTVQTGI	60	
Qy	61	NIAGRIILGVLPVPPAGQIASFYSFLVGLWPRGRDQWEIFLEHVEQLINQQTENARNTA	120	
Db	61	NIAGRIILGVLPVPPAGQIASFYSFLVGLWPRGRDQWEIFLEHVEQLINQQTENARNTA	120	
Qy	121	LARIQGLGDSFRAYQQSLEDWLENRDDARTSVLTQVIALELDFLNAMPLFAIRNQVEVP	180	
Db	121	LARIQGLGDSFRAYQQSLEDWLENRDDARTSVLTQVIALELDFLNAMPLFAIRNQVEVP	180	
Qy	181	LLMWYAQANLHLLLRDASLFGSEFGTISOIRYRYEROVEQTRDYSYCVWNTGLN	240	
Db	181	LLMWYAQANLHLLLRDASLFGSEFGTISOIRYRYEROVEQTRDYSYCVWNTGLN	240	
Qy	241	SLRGNTAASWRYNQFRDLTLGLVLDLVALPSPYDTRTPINTSAQLTREVYVTDIAIGATG	300	
Db	241	SLRGNTAASWRYNQFRDLTLGLVLDLVALPSPYDTRTPINTSAQLTREVYVTDIAIGATG	300	
Qy	301	VNMASMNWYNNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHTIQ	360	
Db	301	VNMASMNWYNNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHTIQ	360	

```
QY 361 SRPIGGLNTSTHGSTNTSINPRLSPFRSDVWVTESVAGVLLMGVLEPIHGVTVPFN 420
Db 361 SRPIGGLNTSTHGSTNTSINPRLSPFRSDVWVTESVAGVLLMGVLEPIHGVTVPFN 420
QY 421 FRNPQTFERGNTANSQPYESQGLQKQSETLPPETTERPNYESVSHRSHIGLISQR 480
Db 421 FRNPQTFERGNTANSQPYESQGLQKQSETLPPETTERPNYESVSHRSHIGLISQR 480
QY 481 VHVVPVSWTHRGAADRTNTSSDSITQIPLVKSFNLSNGTSVSGPGFTGGDIIRTNVNGS 540
Db 481 VHVVPVSWTHRGAADRTNTSSDSITQIPLVKSFNLSNGTSVSGPGFTGGDIIRTNVNGS 540
QY 541 VLSMGLNFNTSLQRYRVRVRAASQTMVLRVTWVGSTTFQCGPSTMSANESLTSQFR 600
Db 541 VLSMGLNFNTSLQRYRVRVRAASQTMVLRVTWVGSTTFQCGPSTMSANESLTSQFR 600
QY 601 FAEPFVGISASGQTAGISISNNAQRTFFHFKIEFIPITATFAEYDLERAQAVNALP 660
Db 601 FAEPFVGISASGQTAGISISNNAQRTFFHFKIEFIPITATFAEYDLERAQAVNALP 660
QY 661 TWTNPRRLKTDVTDHIDOVSNLACLSDERCLDEKRELLKVKYAKRLSDERNLLQDPN 720
Db 661 TWTNPRRLKTDVTDHIDOVSNLACLSDERCLDEKRELLKVKYAKRLSDERNLLQDPN 720
QY 721 FTSINKQPDFISTNEQSNFTSHEQSEHGWGSENITIQEGNDVFKENYVTLPGTFNECY 780
Db 721 FTSINKQPDFISTNEQSNFTSHEQSEHGWGSENITIQEGNDVFKENYVTLPGTFNECY 780
QY 781 PTYLKYKIGESLKYATRYQIRGYIEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
Db 781 PTYLKYKIGESLKYATRYQIRGYIEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
QY 841 PIGRCGEPNRCAPFENWPDLDCCRDGCKCAHSHHFSLDIDIGCTDLHENLGVWVVPK 900
Db 841 PIGRCGEPNRCAPFENWPDLDCCRDGCKCAHSHHFSLDIDIGCTDLHENLGVWVVPK 900
QY 901 IKTOEGHARLGNLFEBEKLPLGALSRVKAEEKWRDKREKLQLETKRVYVTEAKEAVDA 960
Db 901 IKTOEGHARLGNLFEBEKLPLGALSRVKAEEKWRDKREKLQLETKRVYVTEAKEAVDA 960
QY 961 LFVDSQVNLQADTNIGMHAADKLVRIRIAYLSLSVPGVNAEIFELEGRIITATS 1020
Db 961 LFVDSQVNLQADTNIGMHAADKLVRIRIAYLSLSVPGVNAEIFELEGRIITATS 1020
QY 1021 LYDARNVYKNGDFNNGLACWNKGVHDVQOSHRSVLVIPWEAEVSAQVRVCPGRGYIL 1080
Db 1021 LYDARNVYKNGDFNNGLACWNKGVHDVQOSHRSVLVIPWEAEVSAQVRVCPGRGYIL 1080
QY 1081 RVTAYKSGYGGCVTHIEIENNTDELKFKNCBEEVYPTDTGTCNDYTAHQGTAVCNERN 1140
Db 1081 RVTAYKSGYGGCVTHIEIENNTDELKFKNCBEEVYPTDTGTCNDYTAHQGTAVCNERN 1140
QY 1141 AGYEDAYEDVTASVNYKPYTEBETVTDVRRDNHCEYDRGVVNPPLPAGYMTKELEYEP 1200
Db 1141 AGYEDAYEDVTASVNYKPYTEBETVTDVRRDNHCEYDRGVVNPPLPAGYMTKELEYEP 1200
QY 1201 ETDKWIIEIGETGTFIVDSVLELLMEE 1228
Db 1201 ETDKWIIEIGETGTFIVDSVLELLMEE 1228
```

RESULT 2

```
US-08-448-170-8
; Sequence 8, Application US/08448170
; Patent No. 5723758
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Steiman, Steve
; TITLE OF INVENTION: No. 5723758el Bacillus thuringiensis Isolate Denoted
```

```
;
; TITLE OF INVENTION: B.t. PSI58C2, Active Against Lepidopteran Peets, and Genes
; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,170
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 102D.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-448-170-8
;
; Query Match 91.5%; Score 5926.5; DB 1; Length 1227;
; Best Local Similarity 91.8%, Pred. No. 0;
; Matches 1129; Conservative 36; Mismatches 60; Indels 5; Gaps 3;
;
; QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
; Db 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
; QY 61 NIAGRIILGVLPVFPAGQATSFYSLVGLWPRGRDQWEIFLEHVEQLINQITENARNTA 120
; Db 61 NIAGRIILGVLPVFPAGQATSFYSLVGLWPRGRDQWEIFLEHVEQLINQITENARNTA 120
; QY 121 LARLQGLGDSFRAYQOQSLDLEWLNRRDDARTSRVLYTVQVIALELDPLNAPLFAIRNQVSP 180
; Db 121 LARLQGLGDSFRAYQOQSLDLEWLNRRDDARTSRVLYTVQVIALELDPLNAPLFAIRNQVSP 180
; QY 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVQTRDYSYCVWYNTGLN 240
; Db 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVQTRDYSYCVWYNTGLN 240
; QY 241 SLRGTNAASWRYNQFRRLDTLGLVDLVALPSPSYDTRTPINTSAQLTREYVTTAIGATG 300
; Db 241 SLRGTNAASWRYNQFRRLDTLGLVDLVALPSPSYDTRTPINTSAQLTREYVTTAIGATG 300
; QY 301 V--NMASNMWNNNAPSFAIETAVIRSPHLLDLFLEQLTIFSTSSRWASATRHMTYWRGHT 358
; Db 301 V--NMASNMWNNNAPSFAIETAVIRSPHLLDLFLEQLTIFSTSSRWASATRHMTYWRGHT 358
; QY 359 IQSRPIGGGLNTSTHGSTNTSINPRLSPFRSDVWVTESVAGVLLMGVLEPIHGVTVPFN 418
; Db 359 IQSRPIGGGLNTSTHGSTNTSINPRLSPFRSDVWVTESVAGVLLMGVLEPIHGVTVPFN 418
```

Db 361 LBSRTIRGSLSTHGNTNTSINPVLTQFTSRDVRTBTFAGINI--LITTPVNGVPMWAR 418
Qy 419 FNFRNPONTFERGTANYSPQYSGPGLQKDSSTELPPTTERPNYESXSHRLSHGLLSQ 478
Db 419 FNRWPLNSL-RGSLYITIGYGVGTQGLPDSSTELPPTTERPNYESXSHRLSNRLISG 477
Qy 479 SRVHPVSWTHRSADRTNTISDSITQIPLVKSNLNSGTSVWSGPGFTGGDIIRTNVN 538
Db 478 NTLRAPVSWTHRSADRTNTISDSITQIPLVKSNLNSGTSVWSGPGFTGGDIIRTNVN 537
Qy 539 GSVLSGMLFNNTSLQRYRVRYAASQTMVLRVTGSGTTTQDQFPSTMSANESLTSQS 598
Db 538 GSVLSGMLFNNTSLQRYRVRYAASQTMVLRVTGSGTTTQDQFPSTMSANESLTSQS 597
Qy 599 FFAFPPVGISASGSGTAGISISNAGROTFFDKIEPIPTATPEAEYDLERAQAVNA 658
Db 598 FFAFPPVGISASGSGTAGISISNAGROTFFDKIEPIPTATPEAEYDLERAQAVNA 657
Qy 659 LFTNTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLKVKYAKRLSDERNLLQD 718
Db 658 LFTNTNPRRLKTDVTDYHIDEVSNLVACLSDEFCLDEKRELLKVKYAKRLSDERNLLQD 717
Qy 719 PNFTSINKQDPFISTNEQSNFTSIHQSEHGWSGSENITIQEGNDVPKENYVTLPGTNE 778
Db 718 PNFTSINKQDPFISTNEQSNFTSIHQSEHGWSGSENITIQEGNDVPKENYVTLPGTNE 777
Qy 779 CYPITYLYOKIGSELKAYTRYQOLRGVIEDSQDLEIYLIRYNKAKHETLDVPGTESVWPLSV 838
Db 778 CYPITYLYOKIGSELKAYTRYQOLRGVIEDSQDLEIYLIRYNKAKHETLDVPGTESVWPLSV 837
Qy 839 ESPICRGCEPNRCAPHFENWPDLCSCRDGKCAHSHHFSLDIDIGCTDLHENLGVWV 898
Db 838 ESPICRGCEPNRCAPHFENWPDLCSCRDGKCAHSHHFSLDIDIGCTDLHENLGVWV 897
Qy 899 FKIKTQEGHARLGNLEFIEBEKPLGALSRLVRRAEKWRDKREKLOLETQRYVITAKAV 958
Db 898 FKIKTQEGHARLGNLEFIEBEKPLGALSRLVRRAEKWRDKREKLOLETQRYVITAKAV 957
Qy 959 DALFVDSQVRLQADTNIGMIAADKLVRHIREAYLSLSVPGVNAEIFELEGRIITA 1018
Db 958 DALFVDSQVRLQADTNIGMIAADKLVRHIREAYLSLSVPGVNAEIFELEGRIITA 1017
Qy 1019 ISLYDARNVKNXGDFNNGLACVNVKGVHDVQVQSHRSVLVIPWEBAEVSQAVRVCPRGY 1078
Db 1018 ISLYDARNVKNXGDFNNGLACVNVKGVHDVQVQSHRSVLVIPWEBAEVSQAVRVCPRGY 1077
Qy 1079 ILRVTAKEGYCEGCVTHIEIENNTDELKFKNCEBEEVYPTDTGTCNDYTAHQGTAVCN 1138
Db 1078 ILRVTAKEGYCEGCVTHIEIENNTDELKFKNCEBEEVYPTDTGTCNDYTAHQGTAVCN 1137
Qy 1139 RNAGYEDAYEVDTTASVNVKPYEETVTDVRRDNHCEYDRGVNVNPPVPGVMTKELEY 1198
Db 1138 RNAGYEDAYEVDTTASVNVKPYEETVTDVRRDNHCEYDRGVNVNPPVPGVMTKELEY 1197
Qy 1199 PPTDKVWIEIGETEGKPIVDSVELLMEE 1228
Db 1198 PPTDKVWIEIGETEGKPIVDSVELLMEE 1227

RESULT 3
US-08-961-803-9
; Sequence 9, Application US/08961803
; Patent No. 6150589
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stelman, Steve
; TITLE OF INVENTION: No. 6150589el Bacillus thuringiensis Isolate Denoted
; TITLE OF INVENTION: B.t. F5158C2, Active Against Lepidopteran Pests, and Genes
; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jay M. Sanders
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.803
; FILING DATE: 31-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/448,170
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: M/S 102DCD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-961-803-9

Query Match 91.5%; Score 5926.5; DB 3; Length 1227;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1129; Conservative 36; Mismatches 60; Indels 5; Gaps 3;

Qy 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNIDPFVSASTVQTGI 60
Qy 61 NIAGRIILGVLPFAGQIASFYSLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120
Db 61 NIAGRIILGVLPFAGQIASFYSLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120
Qy 121 LARLQGLGDSFRAYCQSLDLEWLNRRDARTSVLTQVIALELDFLNAMPLFAIRNOEVP 180
Db 121 LARLQGLGDSFRAYCQSLDLEWLNRRDARTSVLTQVIALELDFLNAMPLFAIRNOEVP 180
Qy 181 LLMVYAQANLHLILLRDLASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVWEVNTGLN 240
Db 181 LLMVYAQANLHLILLRDLASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVWEVNTGLN 240
Qy 241 SLRGTNAASVRYNQFRRLDLTLGVLDLVALPSPYDTRYPINTSAQLTRFVYTDIGATG 300
Db 241 NLRGTNAESWLRVYQFRRLDLTLGVLDLVALPSPYDTRYPINTSAQLTRFVYTDIGATG 300
Qy 301 V--NMAASVWVNNAPSESALETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTVWRGHT 358
Db 301 APSGFASVWVNNAPSESALETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTVWRGHT 360
Qy 359 IQSRPIGGGLNTSTHGSTNTSINPVRLSFFSRDVVWVWTSYAGVLLWGIYLPBPIHGVPVTR 418

QY 899 PKIKTOEGHARLGNLEFIEBEKPLGALSVRAEKWKWDKREKLEKQLETKRYVTEAKEAV 958
Db 898 PKIKTOEGHARLGNLEFIEBEKPLGALSVRAEKWKWDKREKLEKQLETKRYVTEAKEAV 957
QY 959 DALFVDSQYRNLQADTNIGMHAADKLVRHIREAYLSLSVTPGVNABIFEELEGRITTA 1018
Db 958 DALFVDSQYRNLQADTNIGMHAADKLVRHIREAYLSLSVTPGVNABIFEELEGRITTA 1017
QY 1019 ISLYDARNVKNKGNDFNGLACWNVKGVHDVQOSHRSVLVIPEWAEVSOAVRVCPCGRGY 1078
Db 1018 ISLYDARNVKNKGNDFNGLACWNVKGVHDVQOSHRSVLVIPEWAEVSOAVRVCPCGRGY 1077
QY 1079 ILRVTAAYKEGYGEGCVTTHIEINNTDELKFNCEBEEVYPTDTGTCNDYTAHQGTAVCNS 1138
Db 1078 ILRVTAAYKEGYGEGCVTTHIEINNTDELKFNCEBEEVYPTDTGTCNDYTAHQGTAVCNS 1137
QY 1139 RNAGYEDAYEVDTTASVNYKPTYEETVTVDRDNHCEYDRGVNYPVPAGYMTKELEY 1198
Db 1138 RNAGYEDAYEVDTTASVNYKPTYEETVTVDRDNHCEYDRGVNYPVPAGYMTKELEY 1197
QY 1199 PFTDKVWIEIGETGKPIVDSVLELLMBE 1228
Db 1198 PFTDKVWIEIGETGKPIVDSVLELLMBE 1227

RESULT 5

US-07-951-715A-7
; Sequence 7, Application US/07951715A
; Patent No. 5625136

GENERAL INFORMATION:

; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/951,715A
; FILING DATE: 25-SEP-1992
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-951-715A-7

Query Match 87.4%; Score 5659.5; DB 1; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

QY 27 MDLSPDARIESLCTAEGNNINPLVSASTVOTGINIAGRIILGVLPVGPAGQIASPVSLV 86
Db 1 MDLSPDARIESLCTAEGNNIDPFVSASTVOTGINIAGRIILGVLPVGPAGQIASPVSLV 60
QY 87 GELWPRGRDQWEIFLEHVEQLINQOITENARNTALARIQGLGDSFRAVQOSLEDWLENRD 146
Db 61 GELWPRGRDQWEIFLEHVEQLINQOITENARNTALARIQGLGDSFRAVQOSLEDWLENRD 120
QY 147 DARTRSVLTYOYIALELDFLNAMPLFAIRNOEVPLLMVYQAANLHLLLDASLFGSEF 206
Db 121 DARTRSVLTYOYIALELDFLNAMPLFAIRNOEVPLLMVYQAANLHLLLDASLFGSEF 180
QY 207 GLTSQEIQRVYERQVOTRSDYCVWYNTGLNSLRCNTNAASVVRVYVQFRDLTLGVLD 266
Db 181 GLTSQEIQRVYERQVOTRSDYCVWYNTGLNSLRCNTNAASVVRVYVQFRDLTLGVLD 240
QY 267 LVALFPSTYDTRTYPINTSAQLTREVVYDAIGATGVNMASMMWYNNNAPSFSAIETAVIRS 326
Db 241 LVALFPSTYDTRTYPINTSAQLTREVVYDAIGATGVNMASMMWYNNNAPSFSAIETAVIRS 300
QY 327 PHLLDFLEQLTIPSTSRWSATRMWYWRGHTIOSRPITGGGLNTSTHGSTNTSINPVRLS 386
Db 301 PHLLDFLEQLTIPSTSRWSATRMWYWRGHTIOSRPITGGGLNTSTHGSTNTSINPVRLS 360
QY 387 PFSDRVVWTSYAGVLLWGIYLEPIHGVPTVRFPNPONTFERGTANYSPYSPGLQL 446
Db 361 PFSDRVVWTSYAGVLLWGIYLEPIHGVPTVRFPNPONTFERGTANYSPYSPGLQL 420
QY 447 KDSETELPPTETTERPNYESYSHRSLSHIGLISQSRVHVVPVYSWTHRSADRTNTISDSTIQ 506
Db 421 KDSETELPPTETTERPNYESYSHRSLSHIGLISQSRVHVVPVYSWTHRSADRTNTIGNRTIQ 480
QY 507 IPLVKSFLNSGTSVWSGPGTGGDIIRTNVNGSVLSMGLNPNNTSLQRYVRVRYAASQ 566
Db 481 IPLVKSFLNSGTSVWSGPGTGGDIIRTNVNGSVLSMGLNPNNTSLQRYVRVRYAASQ 540
QY 567 TMVLRTVVGSTTDFDQGPSTMSANESLTSOSRPAEPFVPGVCSAGSQ-TAGISISNAG 625
Db 541 TMVLRTVVGSTTDFDQGPSTMSANESLTSOSRPAEPFVPGVCSAGSQ-TAGISISNAG 600
QY 626 RQTFHFKIEPIPTATFEAYDLERAQEAVALFTNTNPRRLKTDVYHIDQVSNLVA 685
Db 601 RQTFHFKIEPIPTATFEAYDLERAQEAVALFTNTNPRRLKTDVYHIDQVSNLVA 660
QY 686 CLSDFECLDKRELLEKVKYAKRLSDERNLLQDPNFTSINKQPDFISTNEQSNFTSIHEQ 745
Db 661 CLSDFECLDKRELLEKVKYAKRLSDERNLLQDPNFTSINKQPDFISTNEQSNFTSIHEQ 720
QY 746 SEHGWGSENITIOEGNDVFKENVVTLPGTNECVPTLYOKIGESLKVATRYOLRGYI 805
Db 721 SEHGWGSENITIOEGNDVFKENVVTLPGTNECVPTLYOKIGESLKVATRYOLRGYI 780
QY 806 EDSQDLIELYLRVNAKHETLDVPGTESVWPLSVESPIGRCGEPNRCAPHFENPDLCSC 865
Db 781 EDSQDLIELYLRVNAKHETLDVPGTESVWPLSVESPIGRCGEPNRCAPHFENPDLCSC 840
QY 866 RDGSKCAHSHHPSLIDIGCTDLHENLGVWVVKITQEGHARLGNLEFIEBEKPLGEEA 925

Db 841 RDGEKCAHSHFSLDIDVCTDHNENLGVWVFKIKTOEGHARLGNLEFIBEKPLGSA 900
QY 926 LSRVRAEKWRDKREKQLQETKRVVTEAKEAVDALFVDSQYNRLQADTNIGMHAADKL 985
Db 901 LSRVRAEKWRDKREKQLQETKRVVTEAKEAVDALFVDSQYDRQLQADTNIGMHAADKL 960
QY 986 VHRIRAYISELSVIFGVNAEIEELEGRIITAIISLYDARNVVGKGFNFNGLACVNVKXG 1045
Db 961 VHRIRAYISELSVIFGVNAEIEELEGRIITAIISLYDARNVVGKGFNFNGLTCNVKXG 1020
QY 1046 VDVQSHRSVLVPEWEAEVSQAVRCPGGRGYILRVTAKEGYGEGCVTTHIEINNTDE 1105
Db 1021 VDVQSHRSVLVPEWEAEVSQAVRCPGGRGYILRVTAKEGYGEGCVTTHIEINNTDE 1080
QY 1106 LKFNCEEEVEVPTDGTCTNDYTAHQGT-----VNSNRNAGYEDAYEVDVTTASVNYKPT 1161
Db 1081 LKFNCEEEVEVPTDGTCTNDYTAHQGTAGCADACNSRNAGYEDAYEVDVTTASVNYKPT 1140
QY 1162 EBETVTVDRDNHCEYDRGVNYPPLPAGYMTKELEYPPETDKWIEIETEGEKFIIVDSV 1221
Db 1141 EBETVTVDRDNHCEYDRGVNYPVPPVAGYVTKLEYFPETDVTWIEIETEGEKFIIVDSV 1200
QY 1222 ELLMEE 1228
Db 1201 ELLMEE 1207

RESULT 6

US-08-459-448A-7
; Sequence 7, Application US/08459448A
; Patent No. 5859336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Deesai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,448A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027

; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-448A-7

Query Match 87.4%; Score 5659.5; DB 2; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

QY 27 MDLSPDARIEDSLCIAEGNNINPLYSASTVQTGINIAGRIILGVLPFPAGQIASFYSLV 86
Db 1 MDLSPDARIEDSLCIAEGNNIDPFVSASTVQTGINIAGRIILGVLPFPAGQIASFYSLV 60
QY 87 GELWPRGRDQWEIFLEHVEQLINQITENARNALARLQGLGDSFRAYQOQSLDLENRD 146
Db 61 GELWPRGRDQWEIFLEHVEQLINQITENARNALARLQGLGDSFRAYQOQSLDLENRD 120
QY 147 DARTRSVLTYQIALELDFLNAMPLFAIRNQVPLMLVYQAQANLHLLLRDASLFGSEF 206
Db 121 DARTRSVLTYQIALELDFLNAMPLFAIRNQVPLMLVYQAQANLHLLLRDASLFGSEF 180
QY 207 GLTSQEIQRYRQVEQTRDYSDYCVWYNTGLNSLRTNAASWVRVYQNRDLTLGVLD 266
Db 181 GLTSQEIQRYRQVEQTRDYSDYCVWYNTGLNSLRTNAASWVRVYQNRDLTLGVLD 240
QY 267 LVALPFSYDTRTYPINTSAQLTREVTDAIGATGVNMAWMYNNNAPSFAIETAVIRS 326
Db 241 LVALPFSYDTRTYPINTSAQLTREVTDAIGATGVNMAWMYNNNAPSFAIETAVIRS 300
QY 327 PHLLDFLEQLTIFSTSSRWASATRWYWRGHTIOSRPIGGGLNTSTHGSTNTSINPVRLS 386
Db 301 PHLLDFLEQLTIFSTSSRWASATRWYWRGHTIOSRPIGGGLNTSTHGSTNTSINPVRLS 360
QY 387 PFSRDVYWTESYAGVLLMGVILEPIHGVTPTFRNFPQNTFERGTANYSQPYESPGIQL 446
Db 361 FASRDVYWTESYAGVLLMGVILEPIHGVTPTFRNFPQNTFERGTANYSQPYESPGIQL 420
QY 447 KDSETELPPTTERPNYESYSHRLSHIGLISQSRVHVYVSWTHRSADRTNTISDSITQ 506
Db 421 KDSETELPPTTERPNYESYSHRLSHIGLISQSRVHVYVSWTHRSADRTNTIGPNRITQ 480
QY 507 IPLVKSFNLSGTSVSGPGTGDDIIRTNNVSGVLSMGLNFNNNTSLQRYVRVRYAASQ 566
Db 481 IPMVKASLPQGTTVARGPGTGDDILRRNTGPGPIRVTVNGPLQRYRIGFYASTV 540
QY 567 TMLRVTVGSGSTTFDQGPSTMSANESLTSQSRFPAPFPVIGISASGSQ-TAGTISINNAG 625
Db 541 DFDFVSRGTTVNNFRFLTMNSGDELKYNFRRATFTPTTQIQDITRTISQIQLSG 600
QY 626 RQTFHDKIEFIPITATFEAYDLERAQEAVALFNTNPRRLKTDVTDVHIQVSNLVA 685
Db 601 NGEVYIDKIEIIPVTATFEAYDLERAQEAVALFNTNPRRLKTDVTDVHIQVSNLVA 660
QY 686 CLSDEFCLDEKRELLEKVKYAKLSDBERNLLQDPNFTSINKQPDFISTNEQSNFTSIHEQ 745
Db 661 CLSDEFCLDEKRELLEKVKYAKLSDBERNLLQDPNFTSINKQPDFISTNEQSNFTSIHEQ 720
QY 746 SEHGWMGSENIITIQEGNDVFKENYVTLPGTFNENCYPTLYQKIGESLKYATYQIRGYI 805
Db 721 SEHGWMGSENIITIQEGNDVFKENYVTLPGTFNENCYPTLYQKIGESLKYATYQIRGYI 780
QY 806 EDSQDLEIYLIRYNKAKHETLDVPGTSEVMPLSVESPIGRCEPNRCAPHFEWNPDLDCSC 865

Db 781 EDSQLEIYIRYNAKHETLDVPGTESLWPLSVESPIGRGCBPNRCAPHFENWPDLCSC 840
Qy 866 RDEKCAHSHHFSLDIDGCTDLHENLGVWVVKIKTOEGHARLGNLEPTEKPLGEA 925
Db 841 RDEKCAHSHHFSLDIDVGCTDLHENLGVWVVKIKTOEGHARLGNLEPTEKPLGEA 900
Qy 926 LSRVRAEKWDRKREKLQLETKRYVTEAKEAVDALFVDSQYNRLQADTNIGMHAADKL 985
Db 901 LSRVRAEKWDRKREKLQLETKRYVTEAKEAVDALFVDSQYDRLQADTNIGMHAADKL 960
Qy 986 VRIREAYLSLSVPGVNAEIEFELEGRITTAISLYDARNVVKGNDFNGLACWNVKGH 1045
Db 961 VRIREAYLSLSVPGVNAEIEFELEGRITTAISLYDARNVVKGNDFNGLACWNVKGH 1020
Qy 1046 VDVOQSHRSVLVPEWAEVSOAVRVCPCGRGYILRVTAKEGYGEGCVTHIEISNNTDE 1105
Db 1021 VDVOQSHRSVLVPEWAEVSOAVRVCPCGRGYILRVTAKEGYGEGCVTHIEISNNTDE 1080
Qy 1106 LKFKNCEBEEVYPTDTGTCNDYTAHQGTA---VCNSRNAGYEDAYEVDTTASVNYKPTY 1161
Db 1081 LKFKNCEBEEVYPTDTGTCNDYTAHQGTAQACACNSRNAGYEDAYEVDTTASVNYKPTY 1140
Qy 1162 BEETVDRDNRHNCYDRGVNYPPLPAGYMTKELEYPPETDKWIEIGETGKPIVDSV 1221
Db 1141 BEETVDRDNRHNCYDRGVNYPVPAGYVTKLEYPPETDKWIEIGETGKPIVDSV 1200
Qy 1222 ELLMEE 1228
Db 1201 ELLMEE 1207

RESULT 7

US-08-459-595A-7
; Sequence 7, Application US/08459595A
; Patent No. 6018104

GENERAL INFORMATION:

;; APPLICANT: Koziel, Michael G.
;; APPLICANT: Desai, Nalini M.
;; APPLICANT: Lewis, Kelly S.
;; APPLICANT: Kramer, Vance C.
;; APPLICANT: Warren, Gregory W.
;; APPLICANT: Evola, Stephen V.
;; APPLICANT: Crossland, Lyle D.
;; APPLICANT: Wright, Martha S.
;; APPLICANT: Merlin, Ellis J.
;; APPLICANT: Launis, Karen L.
;; APPLICANT: Rothstein, Steven J.
;; APPLICANT: Bowman, Cindy G.
;; APPLICANT: Dawson, John L.
;; APPLICANT: Dunder, Erik M.
;; APPLICANT: Pace, Gary M.
;; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 601810artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,595A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/951,715
;; FILING DATE: 25-SEP-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/772,027
;; FILING DATE: 04-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Pace, Gary M.
;; REGISTRATION NUMBER: 40403
;; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (919)541-8582
;; TELEFAX: (919)541-8689
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1207 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
; US-08-459-595A-7

Query Match 87.4%; Score 5659.5; DB 3; Length 1207;

Best Local Similarity 89.4%; Pred. No. 0;

Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

Qy 27 MDLSPDARIEDSLCIAEGNNINPLVSASTVGTGINIAGRILGVLPAGQIASFYSLV 86
Db 1 MDLSPDARIEDSLCIAEGNNIDPFVSASTVGTGINIAGRILGVLPAGQIASFYSLV 60
Qy 87 GELWPRGRDQWEI FLEHVEQLINQOITENARNTALARLQGLGDSFRAYQOQSLDWNLD 146
Db 61 GELWPRGRDQWEI FLEHVEQLINQOITENARNTALARLQGLGDSFRAYQOQSLDWNLD 120
Qy 147 DARTSVLYTCYIALELDFLNAMPLFAIRNOEVLPMVYQAANLHLLLDASLFGSEF 206
Db 121 DARTSVLYTCYIALELDFLNAMPLFAIRNOEVLPMVYQAANLHLLLDASLFGSEF 180
Qy 207 GLTSQEIQRVYERQVOTRDYSDYCVWYNTGLNSLRGTNAASWVRYNQFRDLTLGVLD 266
Db 181 GLTSQEIQRVYERQVOTRDYSDYCVWYNTGLNSLRGTNAASWVRYNQFRDLTLGVLD 240
Qy 267 LVALFPSYDTRTYPTINTSAQLTREYTTDAIGATGVNMAWNNNAPSPAIEFAVIRS 326
Db 241 LVALFPSYDTRTYPTINTSAQLTREYTTDAIGATGVNMAWNNNAPSPAIEFAVIRS 300
Qy 327 PHLDLEQLTIFSTSRWSATRWYWRGHTIOSRPITGGGINTSTHGTSWTSINPVRLS 386
Db 301 PHLDLEQLTIFSTSRWSATRWYWRGHTIOSRPITGGGINTSTHGTSWTSINPVRLS 360
Qy 387 FFSRDVYVTSYAGVLLWGIYLEPIHGVPTVRFNFRPNQNTFERGTANYSPQYSPGLQL 446
Db 361 FASRDVYVTSYAGVLLWGIYLEPIHGVPTVRFNFRPNQNTFERGTANYSPQYSPGLQL 420
Qy 447 KQSETLPPETTERPNYESYSHRLSHIGLSQSRVHVVPVSWTHRSADRTNTISDSTIQ 506
Db 421 KQSETLPPETTERPNYESYSHRLSHIGLSQSRVHVVPVSWTHRSADRTNTISDSTIQ 480
Qy 507 IPVYKSNLNGSTSVSGPGTGGDIIRTNVNGSVLSMGLNPNNTSLQRYVRVRYAASQ 566
Db 481 IPVYKSNLNGSTSVSGPGTGGDIIRTNVNGSVLSMGLNPNNTSLQRYVRVRYAASQ 540
Qy 567 TMLRVTVGSGTTFDQGPSTMSANESLTSQSPFAEPFVPGISASGSO-TAGISISNAG 625
Db 541 DFDFVSRGGTNNFRFLRMNSGDELKYGNFVRRAFTTPTTQIQDIIRTSIQGUSG 600
Qy 626 RQTFHFKIEPIPTATFEAYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 685
Db 601 NGEVYIDKIEIIPVTATFEAYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 660
Qy 686 CLSDEFCLDEKRELLEKVKYAKRLSDERNLQDNPFTSINKQDPFISTNEOSNTSHEQ 745
Db 661 CLSDEFCLDEKRELLEKVKYAKRLSDERNLQDNPFTSINKQDPFISTNEOSNTSHEQ 720

Qy 746 SEHCWGSNITIOEGNDVFNKENVYTLPGTFNCEPTYLYKQIGESSELKAYTRYQLRGYI 805
Db |||||
Qy 721 SEHCWGSNITIOEGNDVFNKENVYTLPGTFNCEPTYLYKQIGESSELKAYTRYQLRGYI 780
Db |||||
Qy 806 EDSQLEIYLIRYNAKHETLDVPGTESVWPLSPVESPIGRCPNRCAPHFENWPDLDSC 865
Db |||||
Qy 781 EDSQLEIYLIRYNAKHETLDVPGTESVWPLSPVESPIGRCPNRCAPHFENWPDLDSC 840
Db |||||
Qy 866 RDGEKCAHSHHESLDIDGCTDLHENLGVWVVFVKIKTOEGHARLGNLEFIEBEKPLGEEA 925
Db |||||
Qy 841 RDGEKCAHSHHESLDIDGCTDLHENLGVWVVFVKIKTOEGHARLGNLEFIEBEKPLGEEA 900
Db |||||
Qy 926 LSRVKAERKWRDKREKLEKQLETKRVVTEAKEAVDALFVDSQYRNRLQADTNIGMHAADKL 985
Db |||||
Qy 901 LSRVKAERKWRDKREKLEKQLETKRVVTEAKEAVDALFVDSQYRNRLQADTNIGMHAADKL 960
Db |||||
Qy 986 VHRIRAYLSELVPGVNAEIPBELEGRIITAIISLYDARNVVKNGDFNNGLACWNVKGH 1045
Db |||||
Qy 961 VHRIRAYLSELVPGVNAEIPBELEGRIITAIISLYDARNVVKNGDFNNGLACWNVKGH 1020
Db |||||
Qy 1046 VDVOQSHHESLDVPEWEAEVQAVRVCPCGRGYILRVTAAYKEGEGCVTIEIENNTDE 1105
Db |||||
Qy 1021 VDVOQSHHESLDVPEWEAEVQAVRVCPCGRGYILRVTAAYKEGEGCVTIEIENNTDE 1080
Db |||||
Qy 1106 LKFKCEBEEVPTDTGTCDNDYTAHQGTAGTACACNSRNAGYEDAYEVDVTTASVNYKPTY 1161
Db |||||
Qy 1081 LKFKCEBEEVPTDTGTCDNDYTAHQGTAGTACACNSRNAGYEDAYEVDVTTASVNYKPTY 1140
Db |||||
Qy 1162 EBEYTDVRRDNHCEYDRGYNYPPLPAGYMTKELEYPEDTKVWIEBETEGKFIIVDSV 1221
Db |||||
Qy 1141 EBEYTDVRRDNHCEYDRGYNYPPLPAGYMTKELEYPEDTKVWIEBETEGKFIIVDSV 1200
Db |||||
Qy 1222 ELLLMEE 1228
Db |||||
Qy 1201 ELLLMEE 1207
Db |||||

RESULT 8

US-08-459-504B-7

; Sequence 7, Application US/08459504B

; Patent No. 6075185

; GENERAL INFORMATION:

; APPLICANT: Koziel, Michael G.

; APPLICANT: Desai, Nalini M.

; APPLICANT: Lewis, Kelly S.

; APPLICANT: Kramer, Vance C.

; APPLICANT: Warren, Gregory W.

; APPLICANT: Evola, Stephen V.

; APPLICANT: Crossland, Lyle D.

; APPLICANT: Wright, Martha S.

; APPLICANT: Merlin, Ellis J.

; APPLICANT: Launis, Karen L.

; APPLICANT: Rothstein, Steven J.

; APPLICANT: Bowman, Cindy G.

; APPLICANT: Dawson, John L.

; APPLICANT: Dunder, Erik M.

; APPLICANT: Pace, Gary M.

; APPLICANT: Suttie, Janet L.

; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED

; NUMBER OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE

; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6075185artis Corporation

; STREET: 3054 Cornwallis Road

; CITY: Research Triangle Park

; STATE: NC

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,504B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-504B-7

Query Match 87.4%; Score 5659.5; DB 3; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;
Qy 27 MDLSPIARIEDSLCTAEGNNINPLVSASTVGTGINIAGRIILGVLPVPPAGQIASFYSFLV 86
Db |||||
Qy 1 MDLSPARIEDSLCTAEGNNIDPFVSASTVGTGINIAGRIILGVLPVPPAGQIASFYSFLV 60
Db |||||
Qy 87 GELWPRGRDQWEIPLHVEQLINQOITENARNTALRLQGLGDSFRAYQOSLEWLENRD 146
Db |||||
Qy 61 GELWPRGRDQWEIPLHVEQLINQOITENARNTALRLQGLGDSFRAYQOSLEWLENRD 120
Db |||||
Qy 147 DARTSRVLYTQYIALELDPLNAMPLFAIRNOEVLPMVYQAANLHLLLRDASLFGSEF 206
Db |||||
Qy 121 DARTSRVLYTQYIALELDPLNAMPLFAIRNOEVLPMVYQAANLHLLLRDASLFGSEF 180
Db |||||
Qy 207 GLTSQEIORYYERQVETRDYSDYCVWYNTGLNSLRGTNAASWVRYNQFRDITLGVLD 266
Db |||||
Qy 181 GLTSQEIORYYERQVETRDYSDYCVWYNTGLNSLRGTNAASWVRYNQFRDITLGVLD 240
Db |||||
Qy 267 LVALFPSYDTRTYPIINTSAQLTREVVYTDALGATGVNMAWNNYNNAPSAIETAVIRS 326
Db |||||
Qy 241 LVALFPSYDTRTYPIINTSAQLTREVVYTDALGATGVNMAWNNYNNAPSAIETAVIRS 300
Db |||||
Qy 327 PHLLDFLEQLTIFSTSRWSATHMTYWRGHTTQSRPIGGGLNTSTHGNTSINPVRLS 386
Db |||||
Qy 301 PHLLDFLEQLTIFSTSRWSATHMTYWRGHTTQSRPIGGGLNTSTHGNTSINPVRLS 360
Db |||||
Qy 387 FFSRDVYWTESYAGVLLWGIYLEPIHGVPTVFRNFRNPQNTFERGTANTYQSPESGLQL 446
Db |||||
Qy 361 FASRDVYWTESYAGVLLWGIYLEPIHGVPTVFRNFRNPQNTFERGTANTYQSPESGLQL 420
Db |||||
Qy 447 KDSLETLPPTTERPNVYESYHRLSHLIGLSQSRVHVYVSWTHRSADRNTTSSDSITQ 506
Db |||||
Qy 421 KDSLETLPPTTERPNVYESYHRLSHLIGLSQSRVHVYVSWTHRSADRNTTSSDSITQ 480
Db |||||
Qy 507 IPLVKSFNLSNGTSVVGPGFTGDDIIRTNVNGSVLSMGLNFNTSLQRYRVRYVAAQ 566
Db |||||
Qy 481 IPWKASLPGQTVVVRGPGFTGDDIIRTNVNGSVLSMGLNFNTSLQRYRVRYVAAQ 540
Db |||||
Qy 567 TMVLRYTVGGSTTFDQGPSTMSANESLTSQSFRAEFVFGISASGSQ-TAGISISNNAG 625
Db |||||
Qy 541 DFDFFVSRGGTIVNNFRFLRTMNSGDELKYGFRVRAFTTPTFTQIQDIIRTSIQGLSG 600
Db |||||
Qy 626 RQTFHFDKIEFIPITATFEAYDLERAQEAVALFTNTNPRILKTDVTDYHIDQVNLVA 685
Db |||||

Db 601 NGEVVIDKIBIIPVATFEAYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVA 660
Qy 686 CLSDFCLDEKELLEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQ 745
Db 661 CLSDFCLDEKELLEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQ 720
Qy 746 SHGWGSGENITIQBNDVFNKYNVTLPGTFNECYPTLYQKIGSELKAYTRYQLRGYI 805
Db 721 SHGWGSGENITIQBNDVFNKYNVTLPGTFNECYPTLYQKIGSELKAYTRYQLRGYI 780
Qy 806 EDSQLELYIRYNKHEITLDVPGTESVWPLSVESPIGRCEPNRCAPHFENPDLDCSC 865
Db 781 EDSQLELYIRYNKHEITLDVPGTESVWPLSVESPIGRCEPNRCAPHFENPDLDCSC 840
Qy 866 RDEKCAHSHHFSLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLEFTEKPLLGEA 925
Db 841 RDEKCAHSHHFSLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLEFTEKPLLGEA 900
Qy 926 LSRVKAERKWDKREKQLETKRVYVTEAKEAVDALFVDSQYDRLOADTNIGMHAADKL 985
Db 901 LSRVKAERKWDKREKQLETKRVYVTEAKEAVDALFVDSQYDRLOADTNIGMHAADKL 960
Qy 986 VHRIRAYLSLSVTPGVNAEIFELEGGHIIITAIISLYDARNVVKNGDFNGLTCWNVKXG 1045
Db 961 VHRIRAYLSLSVTPGVNAEIFELEGGHIIITAIISLYDARNVVKNGDFNGLTCWNVKXG 1020
Qy 1046 VDVOQSHRSVLVPEWAEVSAQVRVCPGRCGYILRTAYKEGCGCTVTHEIENNTDE 1105
Db 1021 VDVOQSHRSVLVPEWAEVSAQVRVCPGRCGYILRTAYKEGCGCTVTHEIENNTDE 1080
Qy 1106 LKPKNREBEVYPTDGTGTCNDYTAHQGTA---VCNSRNAGYEDAYEDVDTTASVNYKPTY 1161
Db 1081 LKPKNREBEVYPTDGTGTCNDYTAHQGTA---VCNSRNAGYEDAYEDVDTTASVNYKPTY 1140
Qy 1162 BEETTYDVRDNHCEYDRGYNPPPLPAGYMTKELEYFETDKWIBETGEGKFIYDSV 1221
Db 1141 BEETTYDVRDNHCEYDRGYNPPVPVAGYVTVKELEYFETDKWIBETGEGKFIYDSV 1200
Qy 1222 ELLMEE 1228
Db 1201 ELLMEE 1207

RESULT 9
US-08-459-444-7

; Sequence 7, Application US/08459444A
; Patent No. 6121014

; GENERAL INFORMATION:

; APPLICANT: Koziel, Michael G.

; Desai, Nalini M.

; Lewis, Kelly S.

; Kramer, Vance C.

; Warren, Gregory W.

; Evola, Stephen V.

; Crossland, Lyle D.

; Wright, Martha S.

; Merlin, Ellis J.

; Launig, Karen L.

; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED

; NUCLEIC ACID CODING SEQUENCE

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.

; STREET: 3054 Cornwallis Road

; CITY: Research Triangle Park

; STATE: NC

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,444A
; FILING DATE: 02-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805/PI/CGC1577/CIP/DIV6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-459-444-7

Query Match 87.4%; Score 5659.5; DB 3; Length 1207;

Beat Local Similarity 89.4%; Pred. No. 0;

Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

Qy 27 MDLSPDARIEDSLCIAEGNNINPLVSASTVGTGINIAGRILGVLPAGQIASPYSFLV 86
Db 1 MDLSPDARIEDSLCIAEGNNIDPPVSASTVGTGINIAGRILGVLPAGQIASPYSFLV 60
Qy 87 GELPRGDDQWEI FLEHVEQLINQOITENARNTALARLQGLGDSFRAYQOQSLDWNLRD 146
Db 61 GELPRGDDQWEI FLEHVEQLINQOITENARNTALARLQGLGDSFRAYQOQSLDWNLRD 120
Qy 147 DARTRSVLTYQYIALELDFLNAMPFAIRNOEVPLLMVYAOANLHLLLRDASLFGSEF 206
Db 121 DARTRSVLTYQYIALELDFLNAMPFAIRNOEVPLLMVYAOANLHLLLRDASLFGSEF 180
Qy 207 GLTSQEIQRYYERQVQTRDYSYCVIEWYNTGLSLRGNTAASWVRYNQFRDLTLGVLD 266
Db 181 GLTSQEIQRYYERQVQTRDYSYCVIEWYNTGLSLRGNTAASWVRYNQFRDLTLGVLD 240
Qy 267 LVALFPSYDRTTYPIINTSAQLTRVYTDATGATGVNMAWNNNAPSPSAIETAVIRS 326
Db 241 LVALFPSYDRTTYPIINTSAQLTRVYTDATGATGVNMAWNNNAPSPSAIETAVIRS 300
Qy 327 PHLLDFLEQLTIPSTSSRSATRMWYRGHTIQSRPIGGGLNTSTHGNTNTINPVRLS 386
Db 301 PHLLDFLEQLTIPSSSRWSNTRMTWRGHTIQSRPIGGGLNTSTHGNTNTINPVRLS 360
Qy 387 PFSRDVYTESYAGVLLWGIYLEPIHGVTYVRFNRPQNTFPERGTANYSPQSPGLQL 446
Db 361 FASRDVYTESYAGVLLWGIYLEPIHGVTYVRFNRPQNTFPERGTANYSPQSPGLQL 420
Qy 447 KOSTELPPETTERPNYESHRLSHIGLSOSRVHPVYVSWTHRSADRTWTISDSITQ 506
Db 421 KOSTELPPETTERPNYESHRLSHIGLSOSRVHPVYVSWTHRSADRTWTISDSITQ 480
Qy 507 IPLVKSFNLSNGTSVWSGPGTGGDIIRTNVNGSVLSNGLNFNTSLQRVVRVRYAASQ 566
Db 481 IPWKASELPQGTIVTVRGPGTGGDIIRTNVNGSVLSNGLNFNTSLQRVVRVRYAASQ 540
Qy 567 TMLRVTVGGSTTDFDQGPSPSTMSANESLTQSOPFAEPFPVGISAGSQ-TAGISISNNAG 625
Db 541 DFDFFVSRRGGTTVNNFRFLRTMNSGDELKYGNFVRRAPFTTPTTQIQDIIRTSIQGLSG 600
Qy 626 RQTFHFKIEPIPIATATFEAYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 685
Db 601 NGEVVIDKIBIIPVATFEAYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 660

Qy	686	CLSD	EFCLDEKRELL	KKVYAKRLSDERN	LQDPNFTSINK	QDPFISTNEQSNFTSIHQ	745					
Db	661	CLSD	EFCLDEKRELL	KKVYAKRLSDERN	LQDPNFTSINK	QDPFISTNEQSNFTSIHQ	720					
Qy	746	SEHG	WGSENITIQEG	NDVFKENVYTLPG	TNECPTYLYQKIG	ESELKAYTRYQLRGVY	805					
Db	721	SEHG	WGSENITIQEG	NDVFKENVYTLPG	TNECPTYLYQKIG	ESELKAYTRYQLRGVY	780					
Qy	806	EDSQ	DLLEIYLIRY	NAKHETLDVPG	TESVWPLSVES	PIGRGCEPNRCAPHF	WPNPDLDCSC	865				
Db	781	EDSQ	DLLEIYLIRY	NAKHETLDVPG	TESVWPLSVES	PIGRGCEPNRCAPHF	WPNPDLDCSC	840				
Qy	866	RDGE	CAHSHHFS	LDIDIGCTD	LHNLGVWVWF	KIKTQGHARLGNLF	IEEKPLLGEA	925				
Db	841	RDGE	CAHSHHFS	LDIDIGCTD	LHNLGVWVWF	KIKTQGHARLGNLF	IEEKPLLGEA	900				
Qy	926	LSRV	KAEEKWRDK	REKLQLETK	RYTTEAK	EAVDALFVDSQ	VNRLQADTNIGW	HAADKL	985			
Db	901	LSRV	KAEEKWRDK	REKLQLETK	RYTTEAK	EAVDALFVDSQ	VNRLQADTNIGW	HAADKL	960			
Qy	986	VHRI	REAYLSEL	VPJGVNAE	IPEELGR	ITAI	SLYDARNVVKNG	DFNGLACVNVKXH	1045			
Db	961	VHRI	REAYLSEL	VPJGVNAE	IPEELGR	ITAI	SLYDARNVVKNG	DFNGLACVNVKXH	1020			
Qy	1046	VDVQ	QSHRSVL	VIPEEA	ESQAVR	CPGGRG	ILRV	TAYKGEYGGCV	THEIENNTDE	1105		
Db	1021	VDVQ	QSHRSVL	VIPEEA	ESQAVR	CPGGRG	ILRV	TAYKGEYGGCV	THEIENNTDE	1080		
Qy	1106	LKFK	NCEEEVY	PTDGT	CNDYTAHQ	TAHQTA	-----VCNS	RNAGYEDAY	VEDVTTASVNYKPT	1161		
Db	1081	LKFK	NCEEEVY	PTDGT	CNDYTAHQ	TAHQTA	CA	CNSRNAGYEDAY	VEDVTTASVNYKPT	1140		
Qy	1162	BEET	YTDVRR	DNHC	YDRG	VYVPP	PLPAGV	WTKELEY	FPFETDKW	WIEIGETEGK	PFVDSV	1221
Db	1141	BEET	YTDVRR	DNHC	YDRG	VYVPP	PLPAGV	WTKELEY	FPFETDKW	WIEIGETEGK	PFVDSV	1200
Qy	1222	ELL	MEE	1228								
Db	1201	ELL	MEE	1207								

RESULT 10
US-09-053-549-8
Sequence 8, Application US/09053549
Patent No. 6121521
GENERAL INFORMATION:
APPLICANT: Desai, Nalini
TITLE OF INVENTION: NO. 6121521el Insecticidal Protein and Gene
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 6121521artis Corporation
STREET: 3054 Cornwallis Rd.
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053,549
FILING DATE: 01-APR-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1995
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 8:

Db 901 LSRVRAEKWKREKLEKLETRKRVYTEAKEAVDALFVDSQYDRLOQADTNIGMHAADKL 960
QY 986 VHRIRAYLSLSVPGVNAEIPFELEGRIITAIISLYDARNVVKNGDFNGLACVNVKGH 1045
Db 961 VHRIRAYLSLSPVPGVNAEIPFELEGRIITAIISLYDARNVVKNGDFNGLTCNVKGH 1020
QY 1046 VDVOQSHRSVLVPEWEAEVQAVRVCPCRGYILRVTAAYKEGYGEGCVTIHEIENNTDE 1105
Db 1021 VDVOQSHRSVLVPEWEAEVQAVRVCPCRGYILRVTAAYKEGYGEGCVTIHEIENNTDE 1080
QY 1106 LAFKNCEREEVPTDGTGTCNDYTAHQGT- ---VNSRNAGYEDAYEDVDTTASVNVKPTY 1161
Db 1081 LAFKNCEREEVPTDGTGTCNDYTAHQGTAGCADACNSRNAGYEDAYEDVDTTASVNVKPTY 1140
QY 1162 EBETTYDVRDNHCEYDRGVNYPPLPAGYMTKELEYFPETDKWIEIGETEGKFIIVDSV 1221
Db 1141 EBETTYDVRDNHCEYDRGVNYPVPVAGYVTKLEYFPETDVTWIEIGETEGKFIIVDSV 1200
QY 1222 ELLLMEE 1228
Db 1201 ELLLMEE 1207

RESULT 11

US-09-547-422-7
; Sequence 7, Application US/09547422
; Patent No. 6320100
; GENERAL INFORMATION:
; APPLICANT: Korziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6320100artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/547,422
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-547-422-7

Query Match 87.4%; Score 5659.5; DB 3; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

QY 27 MDLSPDARIEDSLCIAEONNINPLVSASTVQTGINIAGRILGLVGPVAGQIASYPSFLV 86
Db 1 MDLSPDARIEDSLCIAEONNIDPFVSASTVQTGINIAGRILGLVGPVAGQIASYPSFLV 60
QY 87 GELPRGRDOWEIEFLEHVQELINQOITENARNTALARIQGLGDSFRAVQOSLEDLWNRD 146
Db 61 GELPRGRDOWEIEFLEHVQELINQOITENARNTALARIQGLGDSFRAVQOSLEDLWNRD 120
QY 147 DARTSRVLYTQVIALELDFLNAMPLFAIRNQEVPVLLMYAQAANLHLLLRDASLFGSEF 206
Db 121 DARTSRVLYTQVIALELDFLNAMPLFAIRNQEVPVLLMYAQAANLHLLLRDASLFGSEF 180
QY 207 GLTQEIQRYYERQVQTRDYSYCVWYNTGLNSLRGTNAASWVRYNQFRDLTLGVLD 266
Db 181 GLTQEIQRYYERQVQTRDYSYCVWYNTGLNSLRGTNAASWVRYNQFRDLTLGVLD 240
QY 267 LVALPSPYDTRTYPINTSAQLTREYVTDATCATGVNMAWMWYNNAPSFSAIETAVIRS 326
Db 241 LVALPSPYDTRTYPINTSAQLTREYVTDATCATGVNMAWMWYNNAPSFSAIETAAIRS 300
QY 327 PHLLDFLEQLTIFSTSSRWASATRMWYRGHTIOSRPICGGGLNTSTHGSTNTSINPVRLS 386
Db 301 PHLLDFLEQLTIFSTSSRWASATRMWYRGHTIOSRPICGGGLNTSTHGSTNTSINPVRLS 360
QY 387 FFSRDVWYTESYAGVLLWGIYLEPIHGVPVTRFNPQNPTFERTGANTYSPYSPGLQL 446
Db 361 FASRDVYRTESYAGVLLWGIYLEPIHGVPVTRFNPQNPTFERTGANTYSPYSPGLQL 420
QY 447 KDSLETLPETTERPNYESYSHRLSHIGLISQSRVHVVPVYSWTHRSADRTWTISDSTIQ 506
Db 421 KDSLETLPETTERPNYESYSHRLSHIGLISQSRVHVVPVYSWTHRSADRTWTISDSTIQ 480
QY 507 IPLVKSNLNSGTSVWSGPGTGDIIIRTNVNGSVLSMGLNFNNTSLQRYRVRVVYASQ 566
Db 481 IPWKASLPGTQTVVRGPGTGDILRRWTGFGPIRVTVNGPLTQRYIGRYASTV 540
QY 567 TWLIRVTVGSTTFDQGPSTMSANESLTSQSFRFAEPFVGISASGSQ-TAGISISNNAG 625
Db 541 DFDFVSRGGTTVNNFRFLRTMNSGDELKYNFVRRAPTPTFTQIQDIIRTSIQGLSG 600
QY 626 RQTFHFKIEIPITATFEAYDLERAQEAVALPTNPRRLKTDVTDYHIDQVSNLVA 685
Db 601 NGEYIDKIEIIPVTATFEAYDLERAQEAVALPTNPRRLKTDVTDYHIDQVSNLVA 660
QY 686 CLSDFCLDEKRELLEKVKAKRLSDERNLQDPNFTSINKOPDFISNEQSNFTSIHEQ 745
Db 661 CLSDFCLDEKRELLEKVKAKRLSDERNLQDPNFTSINKOPDFISNEQSNFTSIHEQ 720
QY 746 SEHGWSGSENIITIEGNDVFKENYVTLPGTFNECYPTYLYQKIGESLKYATRYQLRGYI 805
Db 721 SEHGWSGSENIITIEGNDVFKENYVTLPGTFNECYPTYLYQKIGESLKYATRYQLRGYI 780
QY 806 EDSQDLIYLIRYNAKHETLDVPGTESVWPLSVESPIGRGCEPNRCAPHPFWNPDLDSCS 865
Db 781 EDSQDLIYLIRYNAKHETLDVPGTESVWPLSVESPIGRGCEPNRCAPHPFWNPDLDSCS 840
QY 866 RDGEKCAHSHHFLSDIDIGCTDLHENLGVWVFKITQOEGHARLGNLEFTEEPPLGEEA 925
Db 841 RDGEKCAHSHHFLSDIDIGCTDLHENLGVWVFKITQOEGHARLGNLEFTEEPPLGEEA 900
QY 926 LSRVRAEKWKREKLEKLETRKRVYTEAKEAVDALFVDSQYDRLOQADTNIGMHAADKL 985

Db 901 LSRVRAEKWRDKREKLEQLETKRVYTEAKEAVDALFVDSQYDRLOADTNIGMHAADKL 960
QY 986 VHRIRAYISELVIPIGVNAEIPPELEGRITTAISLYDARNVYXGDFNNGIACWNVKGH 1045
Db 961 VHRIRAYISELVIPIGVNAEIPPELEGRITTAISLYDARNVYXGDFNNGIACWNVKGH 1020
QY 1046 VDVOQSHRSVLVPIPEWEAEVSOAVRVCPCRGVILRVTAKEGYGEGCVTTHEIENNTDE 1105
Db 1021 VDVOQSHRSVLVPIPEWEAEVSOAVRVCPCRGVILRVTAKEGYGEGCVTTHEIENNTDE 1080
QY 1106 LKPKNCEBEVYPTDGTGTCNDYTAHQGTA----VCNSRNAGYEDAYEVDVTTASVNYKPTY 1161
Db 1081 LKPKNCEBEVYPTDGTGTCNDYTAHQGTAGCADACNSRNAGYEDAYEVDVTTASVNYKPTY 1140
QY 1162 EEETVYDVRDNHCEYDRGVYNNPPLPAGYMTKELEYPPETDKWIEIGETGKPIVDSV 1221
Db 1141 EEETVYDVRDNHCEYDRGVYNNPVPVAGYVTKLEYFPETDTVWIEIGETGKPIVDSV 1200
QY 1222 ELLMEE 1228
Db 1201 ELLMEE 1207

RESULT 12

US-09-988-462-7

; Sequence 7, Application US/09988462

; Patent No. 6720488

; GENERAL INFORMATION:

; APPLICANT: Korziel, Michael G.

; Desai, Nalini M.

; Lewis, Kelly S.

; Kramer, Vance C.

; Warren, Gregory W.

; Evola, Stephen V.

; Crossland, Lyle D.

; Wright, Martha S.

; Merilin, Ellis J.

; Launis, Karen L.

; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED

; INSECTICIDAL ACTIVITY IN MAIZE

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Syngenta Biotechnology, Inc.

; STREET: 3054 Cornwallis Road

; CITY: Research Triangle Park

; STATE: NC

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/988,462

; FILING DATE: 20-No. 6720488-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/547,422

; FILING DATE: 11-APR-2000

; APPLICATION NUMBER: US 08/459,504

; FILING DATE: 02-JUN-1995

; APPLICATION NUMBER: US 07/951,715

; FILING DATE: 25-SEP-1992

; APPLICATION NUMBER: US 07/772,027

; FILING DATE: 04-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Meigs, J. Timothy

; REGISTRATION NUMBER: 38,241

; REFERENCE/DOCKET NUMBER: S-188051

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (919)541-8587

; TELEFAX: (919)541-8689

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-988-462-7

Query Match 87.4%; Score 5659.5; DB 4; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;
QY 27 MDLSPDARIEDSLCIAEGNNINPLVSASTVQTGINIAGRIILGVLPFAGQATASFYSFLV 86
Db 1 MDLLPDARIEDSLCIAEGNNIDPFVSASTVQTGINIAGRIILGVLPFAGQATASFYSFLV 60
QY 87 GELWPRGRDQWEIPLFHVQLINQQTENARTALTARLQGLGDSFRAYQQSLEDWLENRD 146
Db 61 GELWPRGRDQWEIPLFHVQLINQQTENARTALTARLQGLGDSFRAYQQSLEDWLENRD 120
QY 147 DARTSRVLYTOYIALELDFINAMPLFAIRNOEVPLLMVYQAANLHLLLRDASLFGSSEF 206
Db 121 DARTSRVLYTOYIALELDFINAMPLFAIRNOEVPLLMVYQAANLHLLLRDASLFGSSEF 180
QY 207 GLTSQEIQRYERQVETRDYSDYCVWYNTGLNSLRGTNAASWVRYNQFRDLTLGVLD 266
Db 181 GLTSQEIQRYERQVETRDYSDYCVWYNTGLNSLRGTNAASWVRYNQFRDLTLGVLD 240
QY 267 LVALPSPDYTRYPINTSAQLTREYVYDAIGATGVNMAWMYNNAPSFSAIETAVIRS 326
Db 241 LVALPSPDYTRYPINTSAQLTREYVYDAIGATGVNMAWMYNNAPSFSAIETAVIRS 300
QY 327 PHLLDFLEQLTIFSTSSRWASATRHMTYWRGHTIQSRPIGGGLTSTHGTSTNTPVRLS 386
Db 301 PHLLDFLEQLTIFSTSSRWASATRHMTYWRGHTIQSRPIGGGLTSTHGTSTNTPVRLS 360
QY 387 PFSRDVYNTESYAGVLLMGVILEPIHGVTYPRFNPQNTFERGTANYSPYESPGIQL 446
Db 361 FASRDVYNTESYAGVLLMGVILEPIHGVTYPRFNPQNTFERGTANYSPYESPGIQL 420
QY 447 KDSETELPETTERENYESYSHLSHIGLIQSQRVHVVPVYSWTHRSADRTNITSDDSTQ 506
Db 421 KDSETELPETTERENYESYSHLSHIGLIQSQRVHVVPVYSWTHRSADRTNITGNRITQ 480
QY 507 IPLVKSFNLSGTSVVGPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQRYRVRYAAASQ 566
Db 481 IPMVKASELPQGTTVVRGPGFTGGDILRRNTWGTGPGPIRVTVNGPLTQRYRIGFYASTV 540
QY 567 TMLVLRVTGSGSTTPOGFPSTMSANESLTSQSPFAEPFVIGISAGSQ-TAGISISNNAG 625
Db 541 DFDPFVSRGGTIVNNFRFLRTMNSGDELKYNFVRRATFTPTFTQDIIRTSIQGLSG 600
QY 626 RQTFHDKIEFIPITATPEAEYDLERAQEAVALFTNTNPRRLKTDVTDVHIDOVSNLVA 685
Db 601 NGEVVIDKIEIIPVTATPEAEYDLERAQEAVALFTNTNPRRLKTDVTDVHIDOVSNLVA 660
QY 686 CLSDEFCLDEKRELLEKVKANRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIEHQ 745
Db 661 CLSDEFCLDEKRELLEKVKANRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIEHQ 720
QY 746 SEHGWSGSENIITQEGNDVFNKENVVTLPGTFNECYPTLYQKIGESLKYATRYQLRGYI 805
Db 721 SEHGWSGSENIITQEGNDVFNKENVVTLPGTFNECYPTLYQKIGESLKYATRYQLRGYI 780
QY 806 EDSQDLBIYLRVNAKHETLDVPGTESVPLSVESPIGRCGEPNRCAPHFEMWPDLDCCSC 865
Db 781 EDSQDLBIYLRVNAKHETLDVPGTESVPLSVESPIGRCGEPNRCAPHFEMWPDLDCCSC 840
QY 866 RDGEKCAHSHHPSLDIDIGCTDLHENLGYVWVFKITQEGHARLGNLEFIEBKPLLGEA 925
Db 841 RDGEKCAHSHHPSLDIDIGCTDLHENLGYVWVFKITQEGHARLGNLEFIEBKPLLGEA 900

QY 926 LSRVKAERKWRDREKLEQLETKRVVTEAKEAVDALFVDSQYNRLQADTNGIHAADKL 985
DB 901 LSRVKAERKWRDREKLEQLETKRVVTEAKEAVDALFVDSQYDRQADTNGIHAADKL 960
QY 986 VHRIRAYLSELVPIGVNAEIPFEELEGRITTAISLYDARNVVKGD FNGLA CNVKGH 1045
DB 961 VHRIRAYLSELVPIGVNAEIPFEELEGRITTAISLYDARNVVKGD FNGNGLTCWNVKGH 1020
QY 1046 VDVOQSHRSVLVPIPEWAEVSAVRVCPGRGYILRVTAAYKEGYGEGCVTIHEIENNTDE 1105
DB 1021 VDVOQSHRSVLVPIPEWAEVSAVRVCPGRGYILRVTAAYKEGYGEGCVTIHEIENNTDE 1080
QY 1106 LKPKNCEBEVPTDGTGTCNDYTAHQGTA----VCNSRNAGYEDAYEVDVTTASVNYKPTY 1161
DB 1081 LKPKNCEBEVPTDGTGTCNDYTAHQGTAHQGTAGCADCNSRNAGYEDAYEVDVTTASVNYKPTY 1140
QY 1162 ESETYDVRDNHCEYDRGVNVPPLPAGYMKKELEYFPETDKWIEIGETGKPIVDSV 1221
DB 1141 ESETYDVRDNHCEYDRGVNVPVPGYVTKLEYFPETDVMWIEIGETGKPIVDSV 1200
QY 1222 ELLLMEE 1228
DB 1201 ELLLMEE 1207

RESULT 13

US-09-053-549-2
; Sequence 2, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwalis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-053-549-2

Query Match 83.9%; Score 5436.5; DB 3; Length 1227;
Best Local Similarity 84.3%; Pred. No. 0;
Matches 1040; Conservative 63; Mismatches 118; Indels 13; Gaps 5;

QY 1 LTSNRKNEEINLALSIPAVNHSHTQMDLSPDARTEDSLCIAEGNNINPLVSASTVQTGI 60
DB 1 MTSNRKNEEINL-----AVSNHSAQMDLLPDARTEDSLCIAEGNNIDPFVSASTVQTGI 55
QY 61 NIAGRILGLVGPVPAQIASFYSLVGLWPRGRDQWEIFLHVQLINQOITENARNTA 120

DB 56 NIAGRILGLVGPVPAQIASFYSLVGLWPRGRDQWEIFLHVQLINQOITENARNTA 115
QY 121 LARLOGLGDSFRAYQOOSLEDWLENRRDARTSVLTQYIALELDFNAMPFLPAIRNQEV 180
DB 116 LARLOGLGDSFRAYQOOSLEDWLENRRDARTSVLTQYIALELDFNAMPFLPAIRNQEV 175
QY 181 LLWYQAANLHLLLRDASLFGSFGSLTQSOEIQRYEROVTEOIRYSDYCVWNTNGLN 240
DB 176 LLWYQAANLHLLLRDASLFGSFGSLTQSOEIQRYEROVTEOIRYSDYCVWNTNGLN 235
QY 241 SLRGTNAASVRYNQFRDLTLGLVDLVALPSPDYDTRTPYINTSQAQLTREVYTDAGTG 300
DB 236 SLRGTNAASVRYNQFRDLTLGLVDLVALPSPDYDTRTPYINTSQAQLTREVYTDAGTG 295
QY 301 VNMSMWNYNNAASPSAIEAIAIRSPHLLDLEQLTIFSSASSRWSNRHMTYWRGHTIQ 360
DB 296 VNMSMWNYNNAASPSAIEAIAIRSPHLLDLEQLTIFSSASSRWSNRHMTYWRGHTIQ 355
QY 361 SRPIGGGLNTSTHGSTNTSINPVRLSFFSRDVMYTESYAGVLLWGIYLEPIHGVPTVRFN 420
DB 356 SRPIGGGLNTSTHGSTNTSINPVRLSFFSRDVMYTESYAGVLLWGIYLEPIHGVPTVRFN 415
QY 421 FRNPONTFERGTANYSQPYESFGLQKDSSETPELPTETTERPNYESYSHRLSHIGLIQS 480
DB 416 FRNPONTFERGTANYSQPYESFGLQKDSSETPELPTETTERPNYESYSHRLSHIGLIQS 475
QY 481 VHVPUYSWTHSADRTNTISSDITQIPLVKSFNLSGTSVVGSGPFGTGGDIIRTNVNGS 540
DB 476 VHVPUYSWTHSADRTNTISSDITQIPLVKSFNLSGTSVVGSGPFGTGGDIIRTNVNGS 535
QY 541 VLSMGLNNTSLQRYRVRVYAAASQTMVLAVTVGGSTTFDQGFPTSMANESLTSQSFR 600
DB 536 FGPRTVTVNGPLTQRYRIGFYASTVDFDFVSRGTTVNNFRFRTMNSGDELKYGXNFV 595
QY 601 FAEPPVGISASGQT-AGISISNNAGROTFFDKIEFIPITATFAEYDLERAQAVNAL 659
DB 596 RRAFTPTFTQIQNIIRTSIOGLSGNGEVYDKIEIIPVTATPEAYDLERAQAVNAL 655
QY 660 FTNTNPRKLTVDVTHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERLLODP 719
DB 656 FTNTNPRKLTVDVTHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERLLODP 715
QY 720 NFTSINKQDFISTNEQSNFTSIHQSEHGWGSENIITQEGNDVFKENYVTLPTFTNEC 779
DB 716 NFTSINKQDFISTNEQSNFTSIHQSEHGWGSENIITQEGNDVFKENYVTLPTFTNEC 775
QY 780 YPTLYQKIGSELKAYTRYQLRGYIEDSQDLIYLIRYNKAKHETLDVPGTESVWPLSVE 839
DB 776 YPTLYQKIGSELKAYTRYQLRGYIEDSQDLIYLIRYNKAKHETLDVPGTESVWPLSVE 835
QY 840 SPIGRCEPNRCAPHFENPDLDCSDEKCAHSHHSPDLIDIGCTDLHENLGVWVVF 899
DB 836 SPIGRCEPNRCAPHFENPDLDCSDEKCAHSHHSPDLIDIGCTDLHENLGVWVVF 895
QY 900 KIKTOBQGHARLGNLEFTEKEKLLGEALSRVRAEKWRDREKLEQLETKRVVTEAKEAVD 959
DB 896 KIKTOBQGHARLGNLEFTEKEKLLGEALSRVRAEKWRDREKLEQLETKRVVTEAKEAVD 955
QY 960 ALFVDSQYNRLQADTNGIHAADKL VHRIRAYLSELVPIGVNAEIPFEELEGRITTAI 1019
DB 956 ALFVDSQYNRLQADTNGIHAADKL VHRIRAYLSELVPIGVNAEIPFEELEGRITTAI 1015
QY 1020 SLYDARNVVKGD FNGLA CNVKGHVDV--QSHRSVLVPIPEWAEVSAVRVCPGRGY 1078
DB 1016 SLYDARNVVKGD FNGLA CNVKGHVDV--QSHRSVLVPIPEWAEVSAVRVCPGRGY 1075
QY 1079 ILRVTAAYKEGYGEGCVTIHEIENNTDELKFNKCEEEVYPTDGTGTCNDYTA---HQGTA 1134
DB 1076 ILRVTAAYKEGYGEGCVTIHEIENNTDELKFNKCEEEVYPTDGTGTCNDYTA---HQGTA 1134
QY 1135 VCNRSNAGYEDAYEVDVTTASVNYKPTYEEETTYDVRDNHCEYDRGVNVPVPLPAGYMK 1194

Db 1135 -YTSRNRGCGAVESNSVPADYASAYEEKAYTDGRRDNPCESNRGVDYTPLPAGYVTK 1193

QY 1195 ELEYFPEPDKWIEIGITEGKFIIVDSVELLMEE 1228

Db 1194 ELEYFPEPDKWIEIGITEGTFIIVDSVELLMEE 1227

RESULT 14

US-08-100-709-4

; Sequence 4, Application US/08100709

; Patent No. 5322687

; GENERAL INFORMATION:

; APPLICANT: Donovan, William P.

; APPLICANT: Tan, Yiping

; APPLICANT: Jan, Christine S.

; APPLICANT: Gonzalez Jr., Jose M.

; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5

; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.

; ADDRESS: Nadel

; STREET: 1601 Market Street, 36th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/100,709

; FILING DATE: 19930729

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Egolf, Christopher

; REGISTRATION NUMBER: 27633

; REFERENCE/DOCKET NUMBER: 7205-49

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-757-1590

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1229 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-100-709-4

Query Match 80.8%; Score 5237.5; DB 1; Length 1229;

Best Local Similarity 79.9%; Pred. No. 0;

Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

QY 1 LTSNRKNEEINIALSIPAVNSHSTOMDLPDARDIEDSLCIAEGNNINPLVSASTVQTGI 60

Db 1 LTSNRKNEEINIALSIPVNSSTQNLSPDARDIEDSLCVAEENNIDPFVSASTVQTGI 60

QY 61 NIAGRILGLVGPAGQAFYSLVGLWPRGRDQWEIFLEHVEQLINQQITENARNTA 120

Db 61 NIAGRILGLVGPAGQAFYSLVGLWPRGRDQWEIFLEHVEQLINQQITENARNTA 120

QY 121 LARLQGLGDSFRAYQOQSLDWNLRDARSRLVLYTQYTALELDLFLNAPLFAIRNQVEVP 180

Db 121 LARLEGLGRGYSVQQALETWLNDRDARSRLVLYTQYTALELDLFLNAPLFAIRNQVEVP 180

QY 181 LMVYQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVQTRDYSDYCVQWNTGLN 240

Db 181 LMVYQAANLHLLLRDASLFGSEGMASSDVQYQYQIRYTESNHCQWNTGLN 240

QY 241 SLRGTNAASWVRNQFRDRLTLGLDLVALFSDYTRTPYINTSAQLTREVTDAIGATG 300

Db 241 NLRGTNAESWLRNQFRDRLTLGLDLVALFSDYTRTPYINTSAQLTREIVTDPIGRTN 300

QY 301 V--NWASNNWNNAPSFSAIETAVIRSPHLLDPLEQTLTIFSTSSRSWATRHMTYWRGHT 358

Db 301 APSGFASTNNAPSFSAIETAAIFRPPHLLDPLEQTLTISASSRWSSTQHNNYVWGRH 360

QY 359 IQSRPIGGGLNTSTHGST-NTSINPVRLSPFSDRVVWTESYAGVLLWGLIYFIRHGVPTV 417

Db 361 LNFRPIGGTLNTSTOGLTNTSINPVLTQFTSDRVVWTESNAGTNI--LFTTVNGVPMWA 418

QY 418 RFNRPNPONTFERGTANYSQYESPGQLKMDSETLPPETTERPNVYESVSHRLSHGLIS 477

Db 419 RFNFINPQNIYERGATTYSQYQVGLQFSETLPPETTERPNVYESVSHRLSHGLII 478

QY 478 QSRVHVVPVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNV 537

Db 479 GNTLRAPVSWTHRSADRTNTIGPNRITQIPLKALNLSGTVVVGPGFTGGDILRRTN 538

QY 538 NGSVLSMGLNFNNTSLQRYVRVRYAASQTMVLRVTVGGSTTFDQGFPPSMANESLTSQ 597

Db 539 TGTFGDIRLNLNVPLSQRYVRVRYASTTDLQFFTFINGTTVGNFSTRMARGONLEYR 598

QY 598 SFRFAEPFVGISASGSOTAGISISNNAGQTHFDKIEFIPITATPEAPYDLERAQEA 657

Db 599 SFRTAGFSTPFNLNAQSTFTLGAOSFSNQEVYIDRVFVPAEAVTPEAEYDLERAQKAVN 658

QY 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLDEFCLDEKRELEKVKAKRLSDERNLLO 717

Db 659 ALFTSTNPRRLKTDVTDYHIDQVSNVACLDEFCLDEKRELEKVKAKRLSDERNLLO 718

QY 718 DNFTSINKQDPFISTNEQSNTSIHSEHGWGSENITIQEGNDVFKENYVTLPGTFN 777

Db 719 DNFTFISQLSFASIDGQSNFPIINELSEHGWGSAVNTIQEGNDVFKENYVTLPGTFN 778

QY 778 ECVPTLYQKIGESLKAATRYQLRGYIEDSDLEIYLIRYNKAKHETLDVFGTSEWPLS 837

Db 779 ECVPNLYQKIGESLKAATRYQLRGYIEDSDLEIYLIRYNKAKHETLDVFGTSLWPLS 838

QY 838 VESPIGRCEPNRCAPHFENPDLDCSCRDGCKCAHSHHFLSDIDIGCTDLHENLGVWV 897

Db 839 VESPIGRCEPNRCAPHFENPDLDCSCRDGCKCAHSHHFLTDIDVGTDLHENLGVWV 898

QY 898 VPKIKTQGHARLGNLEFIEEKPLLGEALSRVKAEEKWRDKREKLQLETKRVVTEAKA 957

Db 899 VPKIKTQGYARLGNLEFIEEKPLLGEALSRVKAEEKWRDKREKLQLETKRVVTEAKA 958

QY 958 VDALFVDSQYNRLQADTNIIGHIHAADKLVRIRAYLSLSVIPGVNAEIFEELEGRIT 1017

Db 959 VDALFVDSQYDQLQADTNIIGHIHAADKLVRIRAYLSLSELPVIPGVNAEIFEELEGHIT 1018

QY 1018 AISLYDARNVVKNGDFNNGLACVNVKGVHDVQQSHHRSVLVPEWEAEVSOAVRVCPRG 1077

Db 1019 AWSLYDARNVVKNGDFNNGLTCMVNVKGVHDVQQSHHRSVLVPEWEAEVSOAVRVCPRG 1078

QY 1078 YILRVTAKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGTCNDYTAHQGTAVCN 1137

Db 1079 YILRVTAKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGTCNDYTAHQGTAA 1138

QY 1138 SRNAGYEDAEVDTTASVNYKPYEETTYTDRRDNHCEYDRGVVNYPPPLPAGYMTKELE 1197

Db 1139 SRNAGYEDAEVDTTASVNYKPYEETTYTDRRDNHCEYDRGVVNYPPVAGVYVTELE 1198

QY 1198 YFPETDVTWIEIGTEGKFIIVDSVELLMEE 1228

Db 1199 YFPETDVTWIEIGTEGKFIIVDSVELLMEE 1229

RESULT 15

US-08-176-865-4

; Sequence 4, Application US/08176865

; Patent No. 5616319

; GENERAL INFORMATION:

; APPLICANT: Donovan, William P.

; APPLICANT: Tan, Yiping

APPLICANT: Jany, Christine S.
APPLICANT: Gonzalez Jr., Jose M.
TITLE OF INVENTION: BACILLUS THURINGIENSIS c1yE14 AND c1yE15
TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSES: Panitch Schwarze Jacobs & Nadel c/o A.S.
ADDRESSES: Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,865
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/100,709
FILING DATE: 29-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Egolf, Christopher
REGISTRATION NUMBER: 27633
REFERENCE/DOCKET NUMBER: 7205-49
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1229 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-176-865-4

Query Match 80.8%; Score 5237.5; DB 1; Length 1229;
Best Local Similarity 79.9%; Pred. No. 0;
Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLICVAEGNNINPLVSASTVQTGI 60
DB 1 LTSNRKNEIINALSIPVSNPSQMNLSPDARIEDSLICVAEYNNIDPFVSASTVQTGI 60

QY 61 NIAGRIILGVLPFAGQIASFYSFLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120
DB 61 NIAGRIILGVLPFAGQIASFYSFLVGLWPRGRDQWEIFLEHVQLIRQVNTENTRNTA 120

QY 121 LARLOGLGDSFRAYQGSLEDLWLENRDDARTSVLTVQVIALEDFLNAMPLFAIRNOEVP 180
DB 121 IARLEGLGRGYRQYQQALETWLDNNDARSRIILERYVALELDITTAIPLRINNEEVP 180

QY 181 LLWVYAQAANLHLLRDASLFGSEFGLTSQBIQRYRQVEQTRDYSQVWNTGLN 240
DB 181 LLWVYAQAANLHLLRDASLFGSEWMASSDVNYQEQIRYTESYHCHVQWNTGLN 240

QY 241 SLRGTAASWRYNQFRRLDLTLGVLDLVALPSPYDTRYPINTSAQLTREYVYDAIGATG 300
DB 241 NLRGTAASWRYNQFRRLDLTLGVLDLVALPSPYDTRYPINTSAQLTREYVYDPIGRNT 300

QY 301 V--NNASNNVNNAPSATETAVIRSPHLLDLEQLTIPSTSSRWATRHMTYWRGHT 358
DB 301 APSGFASTWNNNAPSATETAVIRSPHLLDLEQLTIPSTSSRWATRHMTYWRGHT 360

QY 359 IQSRPIGGLANTSTHGST--NTSINPVLSPFSDVYMTESYAGVLLMGWILEPIHGVPTV 417
DB 361 LNFRPIGGLANTSTOGLTNTSINPVLQFSDVYMTESYAGVLLMGWILEPIHGVPTV 418

QY 418 RPNFRNPQNTFERGTANYSQPYESGLQKQSETELPETTERPNYESYSHRLSHIGLIIS 477
DB 418 RPNFRNPQNTFERGTANYSQPYESGLQKQSETELPETTERPNYESYSHRLSHIGLIIS 477

Db 419 RPNFRNPQNTFERGTANYSQPYESGLQKQSETELPETTERPNYESYSHRLSHIGLIIS 478
QY 478 QSRVHVVPVYSWTHRSADRTNTISSDITQIPLVKSFNLSNGTSVVGSGFTGDDIIRTNV 537
Db 479 GNTLRAPVYSWTHRSADRTNTIGPNRITQIPLVKAHLNLSGVTVVGSGFTGDDIIRTNV 538
QY 538 NGSVLSMGLNFNTSLQRYRVRYAASQTMVLRVTVGGSTTFFDQGFPTSMANESLTSQ 597
Db 539 TGTGDIRLNLINPLVSQRVRYRYASTTDLQFFTRINGTTVWNGFSTNRGNLSEYR 598
QY 598 SFRPAEPVPGISASGQTAGISISNNAGROTHFDKIEFIPITATFEAYDIERAQEAVN 657
Db 599 SFRTAGSTFPNPLNAQSTFTLGAQSFNQSVYIDRVFVPAEYTFEAYDIERAQKAVN 658
QY 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLDEFCLEKRELLKVKAKRLSDERNLQ 717
Db 659 ALFTSTNPRRLKTDVTDYHIDQVSNLVACLDEFCLEKRELLKVKAKRLSDERNLQ 718
QY 718 DPNTSINKQDPFISTNEQSNFTSIHQSEHGWGSENITIQEGNDVFKENYVTLPGTFN 777
Db 719 DPNTFISGQLSFASIDQSNFPPSINELSEHGWGSEANVTIQEGNDVFKENYVTLPGTFN 778
QY 778 ECPYTYLYQKIGESSELKAYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDPGTSWPLS 837
Db 779 ECPYNYLYQKIGESSELKAYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDPGTSWPLS 838
QY 838 VESPIGRGCEPNRCAPHPEWNPDLDCSCRDGCKCAHSHHFFSLDIDIGCTDLHENLGVWV 897
Db 839 VESPIGRGCEPNRCAPHPEWNPDLDCSCRDGCKCAHSHHFFSLDIDIGCTDLHENLGVWV 898
QY 898 VFKIQTQEGHARLGNLEFIEBKPLLGALSALSRVKAERKWRDKREKLQLETKRVYTEAKEA 957
Db 899 VFKIQTQEGHARLGNLEFIEBKPLLGALSALSRVKAERKWRDKREKLQLETKRVYTEAKEA 958
QY 958 VDALFVDSQYRQLQADTNIGMIHAADKLVRHRIEAYLSSELVPIGVNABIFEELEGRIT 1017
Db 959 VDALFVDSQYRQLQADTNIGMIHAADKLVRHRIEAYLSSELVPIGVNABIFEELEGRIT 1018
QY 1018 AISLYDARNVVKNGDFNNGLACMNVKGVHDVQOSSHRSVLVIPEWEASVQAVRVCGRG 1077
Db 1019 AMSLYDARNVVKNGDFNNGLTCMNVKGVHDVQOSSHRSVLVIPEWEASVQAVRVCGRG 1078
QY 1078 YILRVATYKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDTGTCNDYTAHQGTAVCN 1137
Db 1079 YILRVATYKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDTGTCNDYTAHQGTAVCN 1138
QY 1138 SRNAGYEDAYEVDVTTASVNYKPTVEETTYTVDRDNCHEYDRGVYVYPPPLPAGYMTKELE 1197
Db 1139 SRNAGYEDAYEVDVTTASVNYKPTVEETTYTVDRDNCHEYDRGVYVYPPPLPAGYMTKELE 1198
QY 1198 YFPETDKVWIEIGETEGKFIVDSVELLIMEE 1228
Db 1199 YFPETDKVWIEIGETEGKFIVDSVELLIMEE 1229

Search completed: November 21, 2004, 17:01:23

Job time : 50 secs

THIS PAGE BLANK (USPTO)

	Query Match	100.0%;	Score 6479;	DB 15;	Length 1228;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	LTSRKNEEIIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSTVQTGI	60		
DB	1	LTSRKNEEIIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSTVQTGI	60		
QY	61	NIAGRILGVLPYFAGQASFSYFLGELWPRGRDQWEIFLHEVQLINQIIVENARNTA	120		

```
Db      61  NIAGRILGVLPFAQIASFYSLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120
Qy      121  LARLOGLGDSFRAYQOQSLDLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNQVEP 180
Db      121  LARLOGLGDSFRAYQOQSLDLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNQVEP 180
Qy      181  LMVYAQAANLHLLLRDASLFGSEFGLTSQBIQRYRQVEQTRDYSQVCVWYNTGLN 240
Db      181  LMVYAQAANLHLLLRDASLFGSEFGLTSQBIQRYRQVEQTRDYSQVCVWYNTGLN 240
Qy      241  SLRGTTNAASWVRYNQPRDLTLGVLDLVALFSPSYDTRTPINTSAQLREVYTDAGTG 300
Db      241  SLRGTTNAASWVRYNQPRDLTLGVLDLVALFSPSYDTRTPINTSAQLREVYTDAGTG 300
Qy      301  VNMAWMNNNAPSAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHTIQ 360
Db      301  VNMAWMNNNAPSAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHTIQ 360
Qy      361  SRPIGGGLNTSTHGSTNTSINPVLSPFSRDVYWTESYAGVLLWGIYLEPIHGVTYVRFN 420
Db      361  SRPIGGGLNTSTHGSTNTSINPVLSPFSRDVYWTESYAGVLLWGIYLEPIHGVTYVRFN 420
Qy      421  FRNPQNTFERGTANYSQPYESPGQLKXDSLTPETTERPNYESYSHRSLHIGLISQSR 480
Db      421  FRNPQNTFERGTANYSQPYESPGQLKXDSLTPETTERPNYESYSHRSLHIGLISQSR 480
Qy      481  VHVPIVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVSGPGTGGDIIRTNVNGS 540
Db      481  VHVPIVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVSGPGTGGDIIRTNVNGS 540
Qy      541  VLSMGLNFNTSLQRYRVRYAASQTMVLRTVYGGSTTFDQGFPTMSANESLTSOSFR 600
Db      541  VLSMGLNFNTSLQRYRVRYAASQTMVLRTVYGGSTTFDQGFPTMSANESLTSOSFR 600
Qy      601  PAEPFVGISAGSQTAGISISNNAAGRTQHPDKIEFIPITATFAEYDLERAQEAVALNF 660
Db      601  PAEPFVGISAGSQTAGISISNNAAGRTQHPDKIEFIPITATFAEYDLERAQEAVALNF 660
Qy      661  TINTPRRLKTDVTDYHIDQVNLVACLSDPECLDEKELLEKKVYAKRLSDERNLQDPN 720
Db      661  TINTPRRLKTDVTDYHIDQVNLVACLSDPECLDEKELLEKKVYAKRLSDERNLQDPN 720
Qy      721  FTSINKQPDFSTNEQSNFTSIHQSEHGMWGSNITIQEGNDVFKENYVTLPGTFNECY 780
Db      721  FTSINKQPDFSTNEQSNFTSIHQSEHGMWGSNITIQEGNDVFKENYVTLPGTFNECY 780
Qy      781  PTYLYQKIGESLKAYTRYQLRGYIEDSDLEIYLIRYNKAKHETLDVPGTGESVWPLSVES 840
Db      781  PTYLYQKIGESLKAYTRYQLRGYIEDSDLEIYLIRYNKAKHETLDVPGTGESVWPLSVES 840
Qy      841  PIGRCGPNRCAPHFENWPNLDCSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
Db      841  PIGRCGPNRCAPHFENWPNLDCSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
Qy      901  IKTOEGHARLGNLEFIEEKLGBALSRVRKAKWRDKREKLQLETKRYVTEAKAVALDA 960
Db      901  IKTOEGHARLGNLEFIEEKLGBALSRVRKAKWRDKREKLQLETKRYVTEAKAVALDA 960
Qy      961  LFVDSQVNRLOADNTNIGMIIHAADKLVRIRIERYLSLSVTPGVNAEIPFELEGRITITAS 1020
Db      961  LFVDSQVNRLOADNTNIGMIIHAADKLVRIRIERYLSLSVTPGVNAEIPFELEGRITITAS 1020
Qy      1021  LYDARNVYKNGDFNNGLACWNVKGHDVQOQSHRSVLVPEWAEVSAQVRCVPCRGYIL 1080
Db      1021  LYDARNVYKNGDFNNGLACWNVKGHDVQOQSHRSVLVPEWAEVSAQVRCVPCRGYIL 1080
Qy      1081  RVTAIYKEGYGEGCVTTHIEIENNTDELKFNCEEEYVPTDTGTCNDYTAHQGTAVCNRSN 1140
Db      1081  RVTAIYKEGYGEGCVTTHIEIENNTDELKFNCEEEYVPTDTGTCNDYTAHQGTAVCNRSN 1140
Qy      1141  AGHEDAYEVDVTTASVNYKPYEBETTYDVRDNHCEYDRGVNYPPLPAGYMTKLEYFP 1200
Db      1141  AGHEDAYEVDVTTASVNYKPYEBETTYDVRDNHCEYDRGVNYPPLPAGYMTKLEYFP 1200
```

```
Db      1141  AGHEDAYEVDVTTASVNYKPYEBETTYDVRDNHCEYDRGVNYPPLPAGYMTKLEYFP 1200
Qy      1201  ETDKQWIEIGETEGKFIVDSEVLLMBE 1228
Db      1201  ETDKQWIEIGETEGKFIVDSEVLLMBE 1228

RESULT 2
US-10-428-961-38
; Sequence 38, Application US/10428961
; Publication No. US20030237111A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE REFERENCE: Polynucleotides, Compositions, and Methods of Use (Amended)
; CURRENT APPLICATION NUMBER: US/10/428,961
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-428-961-38
```

Query Match 99.8%; Score 6464; DB 14; Length 1228;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1224; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60

Qy 61 NIAGRILGVLPFAQIASFYSLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120
Db 61 NIAGRILGVLPFAQIASFYSLVGLWPRGRDQWEIFLEHVQLINQOITENARNTA 120

Qy 121 LARLOGLGDSFRAYQOQSLDLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNQVEP 180
Db 121 LARLOGLGDSFRAYQOQSLDLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNQVEP 180

Qy 181 LMVYAQAANLHLLLRDASLFGSEFGLTSQBIQRYRQVEQTRDYSQVCVWYNTGLN 240
Db 181 LMVYAQAANLHLLLRDASLFGSEFGLTSQBIQRYRQVEQTRDYSQVCVWYNTGLN 240

Qy 241 SLRGTTNAASWVRYNQPRDLTLGVLDLVALFSPSYDTRTPINTSAQLREVYTDAGTG 300
Db 241 SLRGTTNAASWVRYNQPRDLTLGVLDLVALFSPSYDTRTPINTSAQLREVYTDAGTG 300

Qy 301 VNMAWMNNNAPSAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHTIQ 360
Db 301 VNMAWMNNNAPSAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHTIQ 360

Qy 361 SRPIGGGLNTSTHGSTNTSINPVLSPFSRDVYWTESYAGVLLWGIYLEPIHGVTYVRFN 420
Db 361 SRPIGGGLNTSTHGSTNTSINPVLSPFSRDVYWTESYAGVLLWGIYLEPIHGVTYVRFN 420

Qy 421 FRNPQNTFERGTANYSQPYESPGQLKXDSLTPETTERPNYESYSHRSLHIGLISQSR 480
Db 421 FRNPQNTFERGTANYSQPYESPGQLKXDSLTPETTERPNYESYSHRSLHIGLISQSR 480

Qy 481 VHVPIVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVSGPGTGGDIIRTNVNGS 540
Db 481 VHVPIVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVSGPGTGGDIIRTNVNGS 540

QY 541 VLSMGLNFNTSLQRYRVRVRAASQTMVLRVTVGSTTFDQGFPTMSANESLSQSFR 600
Db 541 VLSMGLNFNTSLQRYRVRVRAASQTMVLRVTVGSTTFDQGFPTMSANESLSQSFR 600
QY 601 FPAEPVGISASGQTAGISINNAGROTDFPKIEPIPTATFEAYDLERAQEAVALP 660
Db 601 FPAEPVGISASGQTAGISINNAGROTDFPKIEPIPTATFEAYDLERAQEAVALP 660
QY 661 TWTNPRRLKTDVTDYHIDOVSNLVACLDFECLDEKRELLKVKYAKRLSDERNLLQDPN 720
Db 661 TWTNPRRLKTDVTDYHIDOVSNLVACLDFECLDEKRELLKVKYAKRLSDERNLLQDPN 720
QY 721 FTSINKQPDFISTNEQSNFTSIHQSEHGWGSENIITQEGNDVFKENYVTLPGTFNECY 780
Db 721 FTSINKQPDFISTNEQSNFTSIHQSEHGWGSENIITQEGNDVFKENYVTLPGTFNECY 780
QY 781 PLYLQKIGESLKYATRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
Db 781 PLYLQKIGESLKYATRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
QY 841 PTGRCEPNRCAPHFENWPDLCSDGKCAHSHHFLSLDIDIGCTDLHENLGVVWVFK 900
Db 841 PTGRCEPNRCAPHFENWPDLCSDGKCAHSHHFLSLDIDIGCTDLHENLGVVWVFK 900
QY 901 IKTOEGHARLGNLEFIEEKPLLGEALSRVKRAEKKWRDKREKLQLETKRVYTEAKEAVDA 960
Db 901 IKTOEGHARLGNLEFIEEKPLLGEALSRVKRAEKKWRDKREKLQLETKRVYTEAKEAVDA 960
QY 961 LFDVDSQYNRLQADTNGIMHAADKLVHRIREAYLSLSVPGVNAEIPFEELEGRITTAIS 1020
Db 961 LFDVDSQYNRLQADTNGIMHAADKLVHRIREAYLSLSVPGVNAEIPFEELEGRITTAIS 1020
QY 1021 LDYARNVKNKGFNNGLACWNVKGVHDVQSHRSVVLVPEWAEVSQAVRVCPCRGYTL 1080
Db 1021 LDYARNVKNKGFNNGLACWNVKGVHDVQSHRSVVLVPEWAEVSQAVRVCPCRGYTL 1080
QY 1081 RVTAYKEGEGECVTIHEIENNTDELKPKNCEEEVYPTDCTCNDYTAHQGTAVCNSRN 1140
Db 1081 RVTAYKEGEGECVTIHEIENNTDELKPKNCEEEVYPTDCTCNDYTAHQGTAVCNSRN 1140
QY 1141 AGYDAYEYDVTASVNYKPTVEEYTVDRRNHCEYDRGVYVNPPLPAGYNTKBLEYFP 1200
Db 1141 AGYDAYEYDVTASVNYKPTVEEYTVDRRNHCEYDRGVYVNPPLPAGYNTKBLEYFP 1200
QY 1201 ETDKWIEIGETEGKFIVDSVLLMEE 1228
Db 1201 ETDKWIEIGETEGKFIVDSVLLMEE 1228

RESULT 3
US-10-428-961-63
; Sequence 63, Application US/10428961
; Publication NO. US20030237111A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rugar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; TITLE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amended)
; FILE REFERENCE: MECO201-1
; CURRENT APPLICATION NUMBER: US/10/428,961
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63
; LENGTH: 1227

; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-428-961-63
Query Match 91.3%; Score 5912.5; DB 14; Length 1227;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1127; Conservative 36; Mismatches 62; Indels 5; Gaps 3;
QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 LTSNRKNEIINALSIPAVSNHSAQMLSDARIEDSLCIAEGNNIDPFVSASTVQTGI 60
QY 61 NIAGRILGVLPFAGQIASFYPLVGLWPRGRDQWEIFLEHVHQLINQQTENARNTA 120
Db 61 NIAGRILGVLPFAGQIASFYPLVGLWPRGRDQWEIFLEHVHQLINQQTENARNTA 120
QY 121 LARLOGLGDSFRAYQOOSLEDWLENRDDARTSRVLTQVIALELDFLNAMPLFAIRNQV 180
Db 121 LARLOGLGDSFRAYQOOSLEDWLENRDDARTSRVLTQVIALELDFLNAMPLFAIRNQV 180
QY 181 LLWVYAQANLHLLLRDASLFGSEFGLTSOEIQRYYEROVBOTRDYSDYCVWYNTGLN 240
Db 181 LLWVYAQANLHLLLRDASLFGSEFGLTSOEIQRYYEROVBOTRDYSDYCVWYNTGLN 240
QY 241 SLRGTTAAASVWRYNQFRDLTLGLVDLVALPFSYDTRTYPIINTSAQLTREYVYDAIGATG 300
Db 241 SLRGTTAAASVWRYNQFRDLTLGLVDLVALPFSYDTRTYPIINTSAQLTREYVYDAIGATG 300
QY 301 V--NWASWVWNNAPSPSAIETAVIRSPHLLDFLEQLTIFSTSSRWASATRMWTWRGHT 358
Db 301 APSGFASTWNNAPSPSAIETAVIRSPHLLDFLEQLTIFSTSSRWASATRMWTWRGHT 358
QY 359 IQSRPIGGLANTSTHGNTNTSINPVLRFSSRDVVTWTSYAGVLLGWVLPBIGHVPTVR 418
Db 359 IQSRPIGGLANTSTHGNTNTSINPVLRFSSRDVVTWTSYAGVLLGWVLPBIGHVPTVR 418
QY 419 FNFRNPONTFERGTANYSQPYESPOLQKQDSETELPPTERRPNYSESHRLSHGLISQ 478
Db 419 FNFRNPONTFERGTANYSQPYESPOLQKQDSETELPPTERRPNYSESHRLSHGLISQ 478
QY 479 SRVHVPTVSWTHRSADRTNTSSDSITQIPLVKSFNLSGTSVWSGPGFTGDDIIRTNVN 538
Db 479 SRVHVPTVSWTHRSADRTNTSSDSITQIPLVKSFNLSGTSVWSGPGFTGDDIIRTNVN 538
QY 539 GSVLSMGLNFNTSLQRYRVRVRAASQTMVLRVTVGSTTFDQGFPTMSANESLSQS 598
Db 539 GSVLSMGLNFNTSLQRYRVRVRAASQTMVLRVTVGSTTFDQGFPTMSANESLSQS 598
QY 599 FRPAEPVGISASGQTAGISINNAGROTDFPKIEPIPTATFEAYDLERAQEAVALP 658
Db 599 FRPAEPVGISASGQTAGISINNAGROTDFPKIEPIPTATFEAYDLERAQEAVALP 658
QY 659 LFTNTPRRLKTDVTDYHIDOVSNLVACLDFECLDEKRELLKVKYAKRLSDERNLLQD 718
Db 659 LFTNTPRRLKTDVTDYHIDOVSNLVACLDFECLDEKRELLKVKYAKRLSDERNLLQD 718
QY 719 PNFTSINKQPDFISTNEQSNFTSIHQSEHGWGSENIITQEGNDVFKENYVTLPGTFNE 778
Db 719 PNFTSINKQPDFISTNEQSNFTSIHQSEHGWGSENIITQEGNDVFKENYVTLPGTFNE 778
QY 779 CYPTLYQKIGESLKYATRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
Db 779 CYPTLYQKIGESLKYATRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
QY 839 ESPIGRCGEPNRCAPHFENWPDLCSDGKCAHSHHFLSLDIDIGCTDLHENLGVVWV 898
Db 839 ESPIGRCGEPNRCAPHFENWPDLCSDGKCAHSHHFLSLDIDIGCTDLHENLGVVWV 898
QY 899 FKIKTOEGHARLGNLEFIEEKPLLGEALSRVKRAEKKWRDKREKLQLETKRVYTEAKEAV 958
Db 899 FKIKTOEGHARLGNLEFIEEKPLLGEALSRVKRAEKKWRDKREKLQLETKRVYTEAKEAV 958
QY 959 DALFVDSQYNRLQADTNGIMHAADKLVHRIREAYLSLSVPGVNAEIPFEELEGRITTA 1018

```
Db 958 DALFVDSQYDLQADTNGIMHAADKLVRHIREAYLSLSVPGVNAEIFFEELEGRITTA 1017
Qy 1019 ISLYDARNVVKNGDFNNGLACMNVKGVHDVQOQSHHRSVLVPIPEWEAEVQAARVCPGRGY 1078
Db 1018 ISLYDARNVVKNGDFNNGLACMNVKGVHDVQOQSHHRSVLVPIPEWEAEVQAARVCPGRGY 1077
Qy 1079 ILRTAYKEGEGECVTIHIENNTDELKPKNCBEEVYPTDTGTCDNYTAHQGTAVCNS 1138
Db 1078 ILRTAYKEGEGECVTIHIENNTDELKPKNCBEEVYPTDTGTCDNYTAHQGTAVCNS 1137
Qy 1139 RNAGYEDAYEVDVTTASVNYKPTVEEYTTDVRDNHCEYDRGVVNYPPPLPAGYMTKELEY 1198
Db 1138 RNAGYEDAYEVDVTTASVNYKPTVEEYTTDVRDNHCEYDRGVVNYPPPLPAGYMTKELEY 1197
Qy 1199 FPETDKVWIEBIGETEGKPIVDSVELLMEE 1228
Db 1198 FPETDKVWIEBIGETEGKPIVDSVELLMEE 1227

RESULT 4
US-10-809-953-10
; Sequence 10, Application US/10809953
; Publication No. US20040181825A1
; GENERAL INFORMATION:
; APPLICANT: Van Mellaert, Herman
; APPLICANT: Botterman, Johan
; APPLICANT: Van Rie, Jeroen
; APPLICANT: Joos, Henk
; TITLE OF INVENTION: RECOMBINANT PLANT EXPRESSING NON-COMPETITIVELY BINDING Bt INSECTI
; FILE REFERENCE: 021565-078
; CURRENT APPLICATION NUMBER: US/10/809,953
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US/09/661,016
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: PCT/EP90/00905
; PRIOR FILING DATE: 1990-05-30
; PRIOR APPLICATION NUMBER: GB 89401499.2
; PRIOR FILING DATE: 1989-05-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-809-953-10

Query Match 88.6%; Score 5742; DB 16; Length 1228;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1098; Conservative 35; Mismatches 90; Indels 10; Gaps 3;

Qy 1 LTSRKNEIETINALSPVNSHSTQMDLSPDARIEDSLCIAEGNNINPLVASTVQTGI 60
Db 1 LTSRKNEIETIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVNSTVQTGI 55
Qy 61 NIAGRIILGVLPVGPAGQIASFYLACELWPRGDQWEIFLEHVQIINQITENARNTA 120
Db 56 NIAGRIILGVLPVGPAGQIASFYLWGLWPRGDQWEIFLEHVQIINQITENARNTA 115
Qy 121 LARLQGLGDSFRAYQQSLEDWLENRDDARTSRVLYTQYIALEDLDFLNAMPLFAIRNQEVP 180
Db 116 LARLQGLGDSFRAYQQSLEDWLENRDDARTSRVLYHTQYIALEDLDFLNAMPLFAIRNQEVP 175
Qy 181 LLMVYAAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVQTRDYSYCVWYNTGLN 240
Db 176 LLMVYAAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVQTRDYSYCVWYNTGLN 235
Qy 241 SLRGTNAASVVRNQFRDITLGLVDLVALFPSYDTRTPINTSAQLTRVYTDAGTG 300
Db 236 SLRGTNAASVVRNQFRDITLGLVDLVALFPSYDTRTPINTSAQLTRVYTDAGTG 295
Qy 301 VNMASMNWNNAPSAETAVIRSPHLLDFLEQLTIFSTSSRWSSATRHMTYWRGHTIQ 360
```

```
Db 296 VNMASMNWNNAPSAETAVIRSPHLLDFLEQLTIFSSASSRWSNTRHMTYWRGHTIQ 355
Qy 361 SRPIGGGLNTSTHGSTNTSINPVLSPFSRDVVTWTSYAGVLLWGLYLEPIHGVPVTRFN 420
Db 356 SRPIGGGLNTSTHGATNTSINPVLSPASRDVTRTESIAGVLLWGLYLEPIHGVPVTRFN 415
Qy 421 FRNPQNTFERGTANYSPYSPGLQKQDSELTPEPPTTERPNYESYSHRLSHLGLISQSR 480
Db 416 FTNPQNTSDRGATANYSPYSPGLQKQDSELTPEPPTTERPNYESYSHRLSHLGLISQSR 475
Qy 481 VHVPSVSWTHRSADRTNTSSDSITOIPLVKSPNLASGTSVWVGPGPFTGDIIRTNWNGS 540
Db 476 VNVPSVSWTHRSADRTNTIGPNRITQIPMWKASELPQGTTVVVRGPGFTGCDILRRTNTG 535
Qy 541 VLSMGLNFNNTSLQRYVRVRYAASQTMVLRVTVGGSTTFDQGFSPBMTSANEGLTSQSPR 600
Db 536 FGPIRVTVNGPLTQRYRIGFRYASTVDFDFVSRGGTTVNNFRFLRTMNSGDELKVGNTV 595
Qy 601 FAEFPVVGISASGSQ-TAGISISNNAGROTFFHDKTIEFIPITATFAEYDLERAQAVNAL 659
Db 596 RRAFTTPTFTQIODIIRTSIOGLSNGEVYIDKIEIIPVTATFAEYDLERAQAVNAL 655
Qy 660 FMTNPRRLKTDVTDYHIDQVSNLVACLSDFCLEKRELLKVKYAKRLSDERNLLQDP 719
Db 656 FMTNPRRLKTDVTDYHIDQVSNLVACLSDFCLEKRELLKVKYAKRLSDERNLLQDP 715
Qy 720 NFTSINKQPDFISTNEQSNFTSIHQSEBHGWSNITIQEGNDVFKENYVTLPGTFNEC 779
Db 716 NFTSINKQPDFISTNEQSNFTSIHQSEBHGWSNITIQEGNDVFKENYVTLPGTFNEC 775
Qy 780 YPTLYQKIGESSELKAYTRYQLRGYIEDSQDLLEIYLRYNKAKHETLDVPGTESWPLSVE 839
Db 776 YPTLYQKIGESSELKAYTRYQLRGYIEDSQDLLEIYLRYNKAKHETLDVPGTESWPLSVE 835
Qy 840 SPIGRGCBPNRCAPHFENWPDLDSCRDGKCAHSHHFLSDIDIGCTDLHENLGVWVVF 899
Db 836 SPIGRGCBPNRCAPHFENWPDLDSCRDGKCAHSHHFLSDIDIGCTDLHENLGVWVVF 895
Qy 900 KIKTOEGHARLGNLEFIEBKPLGEALSRRVRAKKWRDKREKLOLETKRVYTEAKEVD 959
Db 896 KIKTOEGHARLGNLEFIEBKPLGEALSRRVRAKKWRDKREKLOLETKRVYTEAKEVD 955
Qy 960 ALFVDSQYDLQADTNGIMHAADKLVRHIREAYLSLSVPGVNAEIFFEELEGRITTA 1019
Db 956 ALFVDSQYDLQADTNGIMHAADKLVRHIREAYLSLSVPGVNAEIFFEELEGRITTA 1015
Qy 1020 SLYDARNVVKNGDFNNGLACMNVKGVHDVQOQSHHRSVLVPIPEWEAEVQAARVCPGRGY 1079
Db 1016 SLYDARNVVKNGDFNNGLTCNVKGVHDVQOQSHHRSVLVPIPEWEAEVQAARVCPGRGY 1075
Qy 1080 LRVTAYKEGEGECVTIHIENNTDELKPKNCBEEVYPTDTGTCDNYTAHQGTAG 1135
Db 1076 LRVTAYKEGEGECVTIHIENNTDELKPKNCBEEVYPTDTGTCDNYTAHQGTAG 1135
Qy 1136 CNSNAGYEDAYEVDVTTASVNYKPTVEEYTTDVRDNHCEYDRGVVNYPPPLPAGYMTKE 1195
Db 1136 CNSNAGYEDAYEVDVTTASVNYKPTVEEYTTDVRDNHCEYDRGVVNYPPPLPAGYMTKE 1195
Qy 1196 LEYPPETDKVWIEBIGETEGKPIVDSVELLMEE 1228
Db 1196 LEYPPETDKVWIEBIGETEGKPIVDSVELLMEE 1228
```

```
RESULT 5
US-09-988-462-7
; Sequence 7, Application US/09988462
; Publication No. US20030046726A1
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
```

Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellie J.
Lauis, Karen L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSER: Syngenta Biotechnology, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/988,462
FILING DATE: 20-No. US20030046726A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/547,422
FILING DATE: 11-APR-2000
APPLICATION NUMBER: US 08/459,504
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-188051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Query Match 87.4%; Score 5659.5; DB 10; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;
QY 27 MDLSPDARIEDSLCAEGNNINPLVSASTVQTGINIAGRILGLVGPVAGQIASFYSFLV 86
DB 1 MDLLPDARIEDSLCAEGNNIDPFVSASTVQTGINIAGRILGLVGPVAGQIASFYSFLV 60
QY 87 GELWPRGRDQWEIFLEHVQLINQIOTENARNTALRLQGLGDSFRAYCQSLDMLNDRD 146
DB 61 GELWPRGRDQWEIFLEHVQLINQIOTENARNTALRLQGLGDSFRAYCQSLDMLNDRD 120
QY 147 DARTSRVLYTQVIALELDFLNAMPFLAIRNQSRVPLLMVYQAANLHLLLRDASLFGSEF 206
DB 121 DARTSRVLYTQVIALELDFLNAMPFLAIRNQSRVPLLMVYQAANLHLLLRDASLFGSEF 180
QY 207 GLTSGEIQRYRQVEQTRDYSQVCEWYNTGLNSLRGTNAASWRYNQFRDLTLGLVLD 266
DB 181 GLTSGEIQRYRQVEQTRDYSQVCEWYNTGLNSLRGTNAASWRYNQFRDLTLGLVLD 240
QY 267 LVALPSPYDTRTPINTSAQLTRVYTDAGATGVMASMMWYNNNAPSFAIETAVIRS 326
DB 241 LVALPSPYDTRTPINTSAQLTRVYTDAGATGVMASMMWYNNNAPSFAIETAVIRS 300

QY 327 PHLLDFLEQLITPSTSSRSWSATRMHTYWRGHTIQSRPIGGGLTSTHGSTNTSINPVRLS 386
DB 301 PHLLDFLEQLITPSSASSRSWSTRMHTYWRGHTIQSRPIGGGLTSTHGATNTSINPVLR 360
QY 387 PFSRDVYMTESYAGVLLWGIYLEPIHGVTYVRFPNPNQPTPERGTANYSQYSPFGLQ 446
DB 361 FASRDVYMTESYAGVLLWGIYLEPIHGVTYVRFPNPNQPTPERGTANYSQYSPFGLQ 420
QY 447 KDSSETLPETTERPNYESYSHRLSHIGLISQSRVHVVPVYSWTHRSADRTWTISDSTQ 506
DB 421 KDSSETLPETTERPNYESYSHRLSHIGLILQSRVNVVPVYSWTHRSADRTWTIGNRITQ 480
QY 507 IPLVKSFNLSGTSVWSGPGTGGDIIRTNVNGSVLSMGLNFNNTSLQRYVRVRYAASQ 566
DB 481 IPWKASSELPGTTVVRGPGTGGDILRRWTGPGPIRVTVNGPLTQRYRIGFYASTV 540
QY 567 TMLRVTVVGGSTTFDQGPSTMSANESLTSQSPFAEPFVGISASGSQ-TAGISISNAG 625
DB 541 DPDFVSRGGTTVANNFRFLRTMWSGDELKYGNFVRAFTTPTFTQIQD1IRTSIQGLSG 600
QY 626 RQTFHFDKIEPIPIATFEAEYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVA 685
DB 601 NGEVYIDKIEIPVATFEAEYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVA 660
QY 686 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPNFTSINKOPDFISTNEQSNFTSIHQ 745
DB 661 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPNFTSINKOPDFISTNEQSNFTSIHQ 720
QY 746 SEHGWMGSENITIOEGNDVFKENYVTLPGTFNECYPTYLYQKIGESSELKAYTRYQLRGYI 805
DB 721 SEHGWMGSENITIOEGNDVFKENYVTLPGTFNECYPTYLYQKIGESSELKAYTRYQLRGYI 780
QY 806 EDSQDLEIYLIRYNAKHETLDVPGTESVWPLSVESPIGRCGEPNRCAPHFENWPLDCSC 865
DB 781 EDSQDLEIYLIRYNAKHETLDVPGTESVWPLSVESPIGRCGEPNRCAPHFENWPLDCSC 840
QY 866 RDGEKCAHSHHFSLDIDIGCTDLHENLGVVWVFKITQEGHARLGNLEFTEEKPLLGBA 925
DB 841 RDGEKCAHSHHFSLDIDIGCTDLHENLGVVWVFKITQEGHARLGNLEFTEEKPLLGBA 900
QY 926 LSRVRAEKKWRDKREKLQLETKRVTYAKBAVDALFVDSQVNRLOADNTNIGMIHAADKL 985
DB 901 LSRVRAEKKWRDKREKLQLETKRVTYAKBAVDALFVDSQVNRLOADNTNIGMIHAADKL 960
QY 986 VHRIREAYLSLSVTPGVNAEIPFELEGRIITATSLYDARNVKNKGFNGLACWNVKGH 1045
DB 961 VHRIREAYLSLSVTPGVNAEIPFELEGRIITATSLYDARNVKNKGFNGLACWNVKGH 1020
QY 1046 VDVOQSHRSVLVTPWEAEVSVQAVRCPGRGYILRVTAKEGEGCVTTHEIENNTDE 1105
DB 1021 VDVOQSHRSVLVTPWEAEVSVQAVRCPGRGYILRVTAKEGEGCVTTHEIENNTDE 1080
QY 1106 LKFKNCEEEVYPTDGTGTCNDYTAHQGTA-----VCNSRNAGYEDAYEDVDTTASVNYKPTY 1161
DB 1081 LKFKNCEEEVYPTDGTGTCNDYTAHQGTAHQGTAHQGTAHQGTAHQGTAHQGTAHQGTA 1140
QY 1162 BEETVTVRRDNHCEYDRGYNYPPPLPAGYNTKELYPETDKWVIEIGTEGKFIIVDSV 1221
DB 1141 BEETVTVRRDNHCEYDRGYNYPPVPPAGYNTKELYPETDKWVIEIGTEGKFIIVDSV 1200
QY 1222 ELLLMEE 1228
DB 1201 ELLLMEE 1207

RESULT 6
US-09-826-660-23
; Sequence 23, Application US/09826660
; Patent No. US20010026940A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.


```
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1186
; TYPE: PR
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Toxin encoded by synthetic B.t. gene
US-09-826-660-23

Query Match      78.8%; Score 5108; DB 9; Length 1186;
Best Local Similarity 80.6%; Pred. NO. 0;
Matches 955; Conservative 76; Mismatches 108; Indels 56; Gaps 8;

QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSDPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MTSNRKNEIINALSIPAVSNHSAQMWLSDARIEDSLCIAEGNNIDPFFVSASTVQTGI 60

QY 61 NIAGRILVLGVPPAGQIASFYLVLGELWPRGRDQWEI|FLEHVEQLINQOITENARNTA 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 NIAGRILVLGVPPAGQIASFYLVLGELWPRGRDPWEI|FLEHVEQLIRQVIENTRDTA 120

QY 121 LARLQGLGDSFRAYQOQSLEDWLNRRDARTSVLYTQYIALELDFLNAMLPFAIRNQVEP 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 LARLQGLGNSFRAYQOQSLEDWLNRRDARTSVLYTQYIALELDFLNAMLPFAIRNQVEP 180

QY 181 LLMVYAQANLHLLLRDASLFGSEFGLTSQEIQRYVERQVETRDYSDYCVWEYNTGLN 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 LLMVYAQANLHLLLRDASLFGSEFGLTSQEIQRYVERQVETREYSDYCARWYNTGLN 240

QY 241 SLRGTTAASVRYNQFRDLTLGVLDLVALPSPVDTRTYPTINTSAQLTRVYVTDATGATG 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 NLRGTTAASVRYNQFRDLTLGVLDLVALPSPVDTRTYPTINTSAQLTRVYVTDPIGRTN 300

QY 301 V--NMASNMVNNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSRWSATRHTMYRGHT 358
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 APSGFASTNFWNNAPSFSAIERAVIRPPHLLDFPEQLTIFSVLSRWSNTQYMYVWVGRH 360

QY 359 IQSRPIGGLNTSTHGSTNTSINPRLSPFSDRYVMTESYAGVLLMGVILEPIHGVPTVR 418
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 LESRTIRGLSTSTHGNTSINPVTIQTSTRDYRTESPAGINI--LLTTPVNGVPMAR 418

QY 419 FNFRNPQTFPRTGTYANYSQVPSGLQKDSFELPPTTERPNYSYSHRLSHIGLISQ 478
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 FNWRNPLNSL-RGSLLYIIGYTGVLQFDSFELPPTTERPNYSYSHRLSNIRLISG 477

QY 479 SRVHPVYVTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVSQFGTGGDIIRTNV 538
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 NTLRAPVYVTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVSQFGTGGDIIRTNV 537

QY 539 GSVLSGLNFNNTSLQRYRVRYAASQTMVLRTVVGSGTTFDQGPPTMSANESLTSQS 598
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
538 GSVLSMGLNFNNTSLQRYRVRYAASQTMVLRTVVGSGTTFDQGPPTMSANESLTSQS 597

QY 599 FRFAEPFVGISASGSGTAGISINNAAGROTFFHDKIEFIPITATPEAYDLERAQEAUNA 658
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
598 FRFAEPFVGISASGSGTAGISINNAAGROTFFHDKIEFIPITATPEASDRLERAQKAVNA 657

QY 659 LPTNTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLEKBEKLEKVKYAKRLSDERNLLQD 718
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
658 LFTSSNQIGLKTVDYTDYHIDRVSNLVECLSDFCLEKBEKLEKVKYAKRLSDERNLLQD 717
```

```
QY 719 PNFTSINKQPDFISTNEQSNFTSIHQSEHGCMWGSSENIITQEGNDVFKENVVTLPGTFNE 778
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
718 PNFRGINRQLD-----RGWGSSTDTITIQGGDDVFKENVVTLGTFDE 759

QY 779 CYPTLYQKIGESSELKAYTRYQURGYIEDSQDLEIYLIRYNAKHETLDVFGTSEVMPLSV 838
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
760 CYPTLYQKIDESKLKAYTRYQURGYIEDSQDLEIYLIRYNAKHETVNVPGTGLWPLSA 819

QY 839 ESPIGRCGEPNRCAPHFWNPDLDCSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWV 898
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
820 PSPIG-----KCAHSHHPSLDDIDVGCTDLNEDLGVWVI 853

QY 899 FKJTOGHARLGNLRFIEBKPLLGALSRVKAERKWRDKREKLOLETKRVVTEAKEAV 958
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
854 FKJTKDGHARLGNLRFIEBKPLVGELARVKAERKWRDKREKLEWEINIVYKEAKESV 913

QY 959 DALFVDSQYNRLQADTNIGMIIHAADKLVHRIIRAYLSLSVIPGVNAEIPFEELEGRITA 1018
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
914 DALFVNSQYDRLOQADTNIAIHAADKRVHSIIRAYLPFLSVIPGVNAAIFFEELEGRIFTA 973

QY 1019 ISLYDARNVYXNGDFNNGLACNVNKGHDV-QQSHHRSVLVVIPEWEAEVSQAVRVCPGRG 1077
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
974 FSLYDARNVIXNGDFNNGLSCWNVKGHDVVEQNNHRSVLVVPWEAEVSQAVRVCPGRG 1033

QY 1078 YILRVTAIYKEGYGEGCVTHIEIENNTDELKFKNCBEEVYPTDTGTCDNYTA----HQGT 1133
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1034 YILRVTAIYKEGYGEGCVTHIEIENNTDELKFKNCBEEVYPTDTGTCDNYTAQEEYEGT 1093

QY 1134 AVCNSNAGYEDAYEDVTTASVNYKPTVEEYTTVDRDNHCEYDRGYVNYPLPAGYMT 1193
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1094 --YTSNRGYDQAYSNSSVPADYASAYEAKYATDGRDRNCPCESNRGYDGYTFLPAGYVT 1151

QY 1194 KELEYFPETDKVMIIBIGETEGKFIVDSVELLMEE 1228
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1152 KELEYFPETDKVMIIBIGETEGTIFVDSVELLMEE 1186

RESULT 7
US-09-972-175-59
; Sequence 59, Application US/09972175
; Publication NO. US20030101482A1
; GENERAL INFORMATION:
; APPLICANT: Baum, Amy Jelen
; Gilmer, James A.
; Mettus, Anne-Marie Light
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
; LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/972,175
; FILING DATE: 05-Oct-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/337,635
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
```


Db 557 TMEIGNLRSFRYTRFDNSPFRANPDIIIGISEQLFGAG-SISSG-----ELYDKIE 611
Qy 636 FIPITATFAEYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDPFCIDE 695
Db 612 IILADATFEASDLERAQAVNALFTSSNQIGLTKTDVTDYHIDQVSNLVACLSDPFCIDE 671
Qy 696 KRELLEKVKYAKRLSDERNLLQDPNPTSINKQDPFISTNEQSNFTSIHQESHGWMGWSN 755
Db 672 KRELSEKVKHAKRLSDERNLLQDPNFRGINRQPD-----RGRGSTD 713
Qy 756 ITIQEGNDVFKENYVTLPGTFNECYTYLYOKIGSELKAYTRYQLRGYIEDSODLEIYL 815
Db 714 ITIQGGDDVFKENYVTLPGTFVDECYTYLYOKIGSELKAYTRYELRGYIEDSODLEIYL 773
Qy 816 IRYNAKHETLDVPGTESVWPLSVESPIGRGCPNRCAPHFENWPNLDSCRCRGECAHHS 875
Db 774 IRYNAKHETLVNPGTGLWPLSAQSPGKCGEPNRCAPHLENNPDLDCSCRCRGECAHHS 833
Qy 876 HPSIDIDIGCTDLHENLGVWVVKIKTQEGHARLGNLEFIEBKPLLGALSRVRAEKK 935
Db 834 HHFTLIDIVGCTDLNEDLGVWVVKIKTQEGHARLGNLEFIEBKPLLGALSRVRAEKK 893
Qy 936 WDKREKOLETKRYVTEAKEAVDALFVDSQYNRLOADTNIGWIIHAADKLVHRIEAYLS 995
Db 894 WDKREKOLETKRYVTEAKEAVDALFVDSQYNRLOADTNIGWIIHAADKLVHRIEAYLS 953
Qy 996 ELSVIPGVNAEIPFEELEGRIITAIISLYDARNVVGNGDFNGLACWNVKGHDV-QQSHR 1054
Db 954 ELSVIPGVNAEIPFEELEGRIITAIISLYDARNVVGNGDFNGLACWNVKGHDV-QQSHR 1013
Qy 1055 SVLVIPWEARVQAVRVCPCRGYILRVTAAYKEGEGCVTTIHEIENNTDELKFCNCEE 1114
Db 1014 SVLVIPWEARVQAVRVCPCRGYILRVTAAYKEGEGCVTTIHEIENNTDELKFCNCEE 1073
Qy 1115 EYVPTDTGTCNDYTA-----HOGTAVCNSENAGYDAYEVDVDTASVNYKPTIEEYTDVR 1170
Db 1074 EYVPTDTGTCNDYTA-----HOGTAVCNSENAGYDAYEVDVDTASVNYKPTIEEYTDVR 1131
Qy 1171 RDNHCYDRGVYVNPPLPAGYVTKLEYPETDKWIEIGETGKPIVDSVELLMBE 1228
Db 1132 RENPCSENRGYVNPPLPAGYVTKLEYPETDKWIEIGETGKPIVDSVELLMBE 1189

RESULT 10
US-09-972-175-2
; Sequence 2, Application US/09972175
; Publication No. US20030101482A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; Gilmer, Amy Jelen
; Mettus, Anne-Marie Light
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
; LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/972,175
; FILING DATE: 05-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/337,635
; FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MRCO:206
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-972-175-2

Query Match 54.0%; Score 3500.5; DB 10; Length 1189;
Best Local Similarity 57.0%; Pred. No. 2.4e-262;
Matches 717; Conservative 143; Mismatches 289; Indels 109; Gaps 20;
Qy 7 NENEII--NALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGINIAG 64
Db 5 NONQCIYPNCLSS-----NPBEVLDDGERISTGN-----SSIDISLVQ 43
Qy 65 RILGVLGVFPFAGQIASFYSLVGLWPRGRDQWEIFLEHVEQLINQOQITENARNTALARL 124
Db 44 FLVSNF-VPGGGFLVGLIDFVWGIVP---SQMDAFLVQIEQLINERAEFARNAIAI 99
Qy 125 QGLGDSFRAYQOQSLDLENDRDARTSVLYTQYITALELDFLNAAMPLFAIRNQEVPLLMV 184
Db 100 EGLGNFIYVEAFKEWEDPNPATRTRVDRFRILDGLLERDIPSEAISFEVPLLSV 159
Qy 185 YQAQANLHLLLRDASLFGSEFGLTSQEIQRYRVEQVQTRDYSDYCVWNTYGLNSLRG 244
Db 160 YQAQANLHLLLRDASLFGSEFGLTSQEIQRYRVEQVQTRDYSDYCVWNTYGLNSLRG 219
Qy 245 TNAASWRYNPFRRDLTLGLVDLVALPSPYDTRTYPINTSAQLTREVYTDATGATGVNNA 304
Db 220 STYQDMITYNLRDLTLTLVDIAAFFPNYDNRYPQIPVQGLTREVYTDPL-----INPN 275
Qy 305 SMNVYNNAPSAIETAVIRSPHLLDLEOLTTPSTSSRSWATRHMTYWRGHTTOSRPI 364
Db 276 PQLOSVAQLPTFNVMESAIKPHPLFDLNLNLTFTD---WFSVGRNFYWGHRVSSLI 332
Qy 365 GGGELTSTHGSTNTSINPVLRSFFSDYVVTESYAGVLL-----WGIYLEPHGVTVRN 420
Db 333 GGGNITSPYIGREANQEPFRSFTFNGPVFRTLNPTLRLLOQPWAPPFNLRGVEGVSEFS 392
Qy 421 FRNPONTPE---RGTYANSQPYESPGLQDKDSELPETTERPNYVESYSHRLSHIGLIS 477
Db 393 --TPNSFTYRGRGTV-----DSLTELPEEDNSVPPREGYSHRLCHATPVQ 436
Qy 478 QSRVHV-----PVYGNWTHRSABRTWTISSDSITQIPLVKSFNLNSGTSVSVGGFTGGDI 533
Db 437 RSGTPEFLTGTGVVFWSTHRSATLTNTIDPERINQIPLVKGFVWGGTSVITGPGFTGGDI 496
Qy 534 RTNVNGSVLSMGLNPNNTSLQRYRVRYAASQ-----TWVLRTVGGSTTTPDQGPSS 586
Db 497 RRNTFGDFVSLQVNNINSPIQRYLRFYASRDRARVIVLGAASTGGVQGVSVNMPLOK 556
Qy 587 TWSANESLTSQSFRAEPP-----VGISASGSQTAGISISNAGRQTFHFDKIE 635
Db 557 TMEIGENLTSFTFYTDSPNPPFRANPDIIIGISEQLFGAG-SISSG-----ELYDKIE 611
Qy 636 FIPITATFAEYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDPFCIDE 695
Db 612 IILADATFEASDLERAQAVNALFTSSNQIGLTKTDVTDYHIDQVSNLVACLSDPFCIDE 671
Qy 696 KRELLEKVKYAKRLSDERNLLQDPNPTSINKQDPFISTNEQSNFTSIHQESHGWMGWSN 755
Db 672 KRELSEKVKHAKRLSDERNLLQDPNFRGINRQPD-----RGRGSTD 713
Qy 756 ITIQEGNDVFKENYVTLPGTFNECYTYLYOKIGSELKAYTRYQLRGYIEDSODLEIYL 815

```
Db 714 ITIQGDDVFKENYVTLPGTVDECPYLYQKIDESKLKAYTRAYELRGVIEDSQDLEIYL 773
Qy 816 IRYNAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENWPDLDSCRDGKCAHHS 875
Db 774 IRYNAKHEIYVNVPTGSLWPLSAQSPGKCGEPNRCAPHLEWNPDLDCSCRDGKCAHHS 833
Qy 876 HHFSLDIDIGCTDLHENLGVVWVFKITQEGHARLGNLFIEBEKPLLGALSRVRAEKK 935
Db 834 HHFTLDIDVGCTDLNEDLGWVIFKJKTQGHARLGNLFIEBEKPLLGALSRVRAEKK 893
Qy 936 WRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRLQADTNIGMIIHAADKLVRHIREAYLS 995
Db 894 WRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRLQADTNIGMIIHAADKLVRHIREAYLS 953
Qy 996 ELSVPGVNAEIPFEELEGRITAIISLYDARNVKNVGNDFNGLACVNVKGVHDV-QQSHHR 1054
Db 954 ELSVPGVNAEIPFEELEGRITAIISLYDARNVKNVGNDFNGLACVNVKGVHDV-QQSHHR 1013
Qy 1055 SVLVIPWEAEVSAVRVCPGRGYILRVYAYKEGYEGECVTIHEIENNTDELKPKNCVEE 1114
Db 1014 SVLVIPWEAEVSAVRVCPGRGYILRVYAYKEGYEGECVTIHEIENNTDELKPKNCVEE 1073
Qy 1115 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTTASVNYKPYESETYTDVR 1170
Db 1074 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTTASVNYKPYESETYTDVR 1131
Qy 1171 RDNHCEYDRGVVNYPPPLPAGYVTKLEYFPETDKVWIEIGTEGKFIVDSVELLMEE 1228
Db 1132 RDNHCEYDRGVVNYPPPLPAGYVTKLEYFPETDKVWIEIGTEGKFIVDSVELLMEE 1189

RESULT 11
US-10-200-522-2
; Sequence 2, Application US/10200522
; Publication No. US20030195336A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Gilmer, Amy Jelen
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE COMPOSITIONS ENCODING LEBIDOPTERAN-T
; FILE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: MECO:213 (11792,0213 DVUS01)
; CURRENT APPLICATION NUMBER: US/10/200,522
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/337,280
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 08/980,071
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 08/757,536
; PRIOR FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Delta Endotoxin
US-10-200-522-2

Query Match 54.0%; Score 3500.5; DB 14; Length 1189;
Best Local Similarity 57.0%; Pred. No. 2,4e-262;
Matches 717; Conservative 143; Mismatches 289; Indels 109; Gaps 20;

Qy 7 NENEII--NALSPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVASASTVQTGTINAG 64
Db 5 NQNCIPYNCLSD-----NPSEVLLDGERISTGN-----SSIDISLSLVQ 43
Qy 65 RILGLGVPPAGIAGFYSFLVGELWPRGRDOWEIEFLEHVEOLINCOITENARNALARL 124
Db 44 FLVSNF-VPGGGFLVGLIDFVWIGVCP---SQWDAFLVQIEQLINERTAEAFARNAAIANL 99
```

```
Qy 125 QGLGDSFRAYQOQSLDMLNRRDDARTSRVLYTQYIALSLDPLNAMPFLFAIRNQEVPLIMV 184
Db 100 EGLGNFNFIYVEAFKEWEEPPNPATRVIRIDFRILDLGLLERDIPSAFISGEVPLLSV 159
Qy 185 YAOAANLHLLLRDASLFGSEFGLTSQEIQRYTERQVEQTRDYSDYCVWVNTGLNSLRG 244
Db 160 YAOAANLHLLAIRDVIFGERWGLTTINVENENRILRIDEVADHCANTYNGLNLPK 219
Qy 245 TNAASWRYNQFPRDLTLGVLDLVALFPPSDTTRTYPIINTSAQLTREVYTDAIGATGVMA 304
Db 220 STYQDMITYNRLRRDLTLVLDIAAFPPYDNNRYPYIQPVGQLTREVYTDDL---INFN 275
Qy 305 SMWYNNNAPSFAIETAVIRSPHLLDFEQLTIFSTSSRSWSATRMTYWRGHTIOSRPI 364
Db 276 PQQSVAQLPTFNWMESSAIRNPHLLFDLNNLTIFTD---WPSVGRNFYWGHRVSSLI 332
Qy 365 GGGINTSTHGSTWTSINPRLSFFSDVDVWTESYAGVLL---WGIYLBPIHGVPTVRFN 420
Db 333 GGGNITSPYIGREANQEPFRSFTFNGPVPFRTLSNPTLRLLQOPWAPPNLRGVEGVFS 392
Qy 421 FRNPONTFE---RGTYANSQPYESPGLQKQDSELTPEPPTERRPNVYESYSHRLSHLGLIS 477
Db 393 --TPTNSFTVRGRGTV-----DSLTELPPEDNSVPPREGYSHRLCHAFVQ 436
Qy 478 QSRVHV----PYVSWTHRSADRTNTISSITQIPLVKSFNLSNGTSVVSFGFTGDI 533
Db 437 RSGTFLTTGTVVFSWTHRSATLNTIDPERINQIPLVKGFRVWGGTSVITFGFTGGDIL 496
Qy 534 RTNVNGSVLSMGLNFNNTSLQRYVRVRYAASQ-----TWVLAVTVGGSTTPOGPPS 586
Db 497 RNTFGDFVSLQVNLNSPITQYRURFYASSRDRARVILVTGAASTGVGGQVSNVMPLOK 556
Qy 587 TMSANESLTSQSFRAEPP-----VGISAGSQTAGISISNNAAGRTFHPDKIE 635
Db 557 TMEIGENLTSRTFRVTFDSNPPSFRANPDIIIGISQPLFGAG-SISGG-----ELYDKIE 611
Qy 636 FIPIATATPEAVDLERAQAVNALFTNTNPRRLKTDVTDYHDIDQVSNLVACLDSDFCLDE 695
Db 612 ILADATPEASDLERAQAVNALFTSSNQIGLTKTDVTDYHDIDQVSNLVACLDSDFCLDE 671
Qy 696 KRELLEKVKYAKRLSDERNLLQDPNFTSINKQPPFISTNEQSNFTSIHQSEHGWGSEN 755
Db 672 KRELSEKVKHAKRLSDERNLLQDPNFRGINQPD-----RGWRGSTD 713
Qy 756 ITIQGDDVFKENYVTLPGTVDECPYLYQKIDESKLKAYTRAYELRGVIEDSQDLEIYL 815
Db 714 ITIQGDDVFKENYVTLPGTVDECPYLYQKIDESKLKAYTRAYELRGVIEDSQDLEIYL 773
Qy 816 IRYNAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENWPDLDSCRDGKCAHHS 875
Db 774 IRYNAKHEIYVNVPTGSLWPLSAQSPGKCGEPNRCAPHLEWNPDLDCSCRDGKCAHHS 833
Qy 876 HHFSLDIDIGCTDLHENLGVVWVFKITQEGHARLGNLFIEBEKPLLGALSRVRAEKK 935
Db 834 HHFTLDIDVGCTDLNEDLGWVIFKJKTQGHARLGNLFIEBEKPLLGALSRVRAEKK 893
Qy 936 WRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRLQADTNIGMIIHAADKLVRHIREAYLS 995
Db 894 WRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRLQADTNIGMIIHAADKLVRHIREAYLS 953
Qy 996 ELSVPGVNAEIPFEELEGRITAIISLYDARNVKNVGNDFNGLACVNVKGVHDV-QQSHHR 1054
Db 954 ELSVPGVNAEIPFEELEGRITAIISLYDARNVKNVGNDFNGLACVNVKGVHDV-QQSHHR 1013
Qy 1055 SVLVIPWEAEVSAVRVCPGRGYILRVYAYKEGYEGECVTIHEIENNTDELKPKNCVEE 1114
Db 1014 SVLVIPWEAEVSAVRVCPGRGYILRVYAYKEGYEGECVTIHEIENNTDELKPKNCVEE 1073
Qy 1115 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTTASVNYKPYESETYTDVR 1170
Db 1074 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTTASVNYKPYESETYTDVR 1131
Qy 1171 RDNHCEYDRGVVNYPPPLPAGYVTKLEYFPETDKVWIEIGTEGKFIVDSVELLMEE 1228
```

Db 1132 RENCESNRGYDTPPLPAGYVTKOLEYFPETDKVWIEIGETEGTIVDSVLELLMEE 1189

RESULT 12

US-10-855-535-2

Sequence 2, Application US/10855535

Publication NO. US20040221334A1

GENERAL INFORMATION:

APPLICANT: Baum, James A.

APPLICANT: Gilmer, Amy Jelen

APPLICANT: Mettuss, Anne-Marie Light

TITLE OF INVENTION: Antibodies Immunoreactive with Lepidopteran-Toxic Polypeptides

FILE REFERENCE: 11792.0214.DVUS02

CURRENT APPLICATION NUMBER: US/10/855,535

CURRENT FILING DATE: 2004-05-27

PRIOR APPLICATION NUMBER: 09/337,635

PRIOR FILING DATE: 1999-06-21

PRIOR APPLICATION NUMBER: 08/980,071

PRIOR FILING DATE: 1997-11-26

PRIOR APPLICATION NUMBER: 08/757,536

PRIOR FILING DATE: 1996-11-27

NUMBER OF SEQ ID NOS: 76

SOFTWARE: PatentIn version 3.2

SEQ ID NO 2

LENGTH: 1189

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Recombinant Delta Endotoxin

US-10-855-535-2

Query Match 54.0%; Score 3500.5; DB 17; Length 1189;

Best Local Similarity 57.0%; Pred. No. 2.4e-262;

Matches 717; Conservative 143; Mismatches 289; Indels 109; Gaps 20;

QY 7 NENEII--NALSIPAVSNHSTQMDLSIPARIBDSLCIAEGNNINPLVBSASTVQTGINAG 64

Db 5 NQNCIPYNCLS-----NPEVLLDGERISTGN-----SSIDISLSLVQ 43

QY 65 RILGVLGVPPAGQIAFSFVLGELWPRGRDOWEIFLEHVEQLINQQTENARNATALARL 124

Db 44 FLVSNF-VPGGFLVLGIDFVGVIGVP---SQWDAFLVQIEQLINERIAEFARNAAIANL 99

QY 125 QSLGDSERAYQOQSLDLENRDDATRSVLYTYQYIALELDPLNAPLEAIRNQEVPLAMV 184

Db 100 EGLGNFNIIYAFKEWEEDPNPATRTVDIFRILDGLLERDIPSPFAISGFEVPLLSV 159

QY 185 YAAANLHLLLRDASLFGSEFGLTSQEIQRYRQVQTRDYSQYCVIEWNTGLNSLRG 244

Db 160 YAAANLHLLLRDASVIFGERGLTTINVENYNRLIRHIDYADHCANTYNRGLNLPK 219

QY 245 TAAASWRYNQPRDLTLGLVDLVALFSDYTRTPINTSAQLTREVTDAIGATGVMA 304

Db 220 STYQDWITYNRLRLDLTLVLDIAAFFNNDNRYPQIPVGQLTREVTDPDLP---INFN 275

QY 305 SNNYNNAPSAETAVIRSPHLLDLEQLTIFSTSSRSATRHMTYRGHTTIQSRPI 364

Db 276 POLQVAQLPTPNVMESSAIRNPHFLDILNNTIFD---WFSVGRNFWGHRVVISLI 332

QY 365 GGGLTSTHGSTNTINPVRISFFSRDVRVWTSYAGVLL---WGYLEPIHGVTVRFN 420

Db 333 CGGNITSPYIGREANQEPSPSTFNGPVFRTLSNPTLRLLQPPAPPNLRGVGVFS 392

QY 421 FRNPQTE---RGYANSQPYESGLQLKQSETELPPETTERPNYSYSHRLSHIGLIS 477

Db 393 --TPTNSFTYRGRTV-----DSLTELPPEDNSVPPREGYSHRLCHATEVQ 436

QY 478 QSRHVH----PVYSWTHRSADRTNTISSDSITQIPLVKSFNLSNSTSVSGPGFTGGDI 533

Db 437 RSGTFLTTGVVFSWTHRSATLTNTIDPERINQIPLVKGFRVWGTSVITGPGFTGGDIL 496

QY 534 RTNVNGSVLSMGLNPNNTSLQRYVRVRYAASQ-----TMVLRTVTGGSTTFQGGFPS 586

Db 497 RENTFGDFVSLQVNLINSPIITQRYRLFRYASSRDARVILTGAASTGVGGQVSNMPLQK 556

QY 587 TMSANESLTSQSRFPABPP-----VGISASGSQTAGISISNNAGRTTFHFDPKTE 635

Db 557 TMEIGENLTSTRTFRYTDSPFPSPFRANPDIIIGISEQPLFGAG-SISSG-----ELYIDKTE 611

QY 636 FIPITATPEAEYDLERAQEAIVNALPTNPRRLKTDVTDYHIDVSNLVACLSDFECLDE 695

Db 612 IILADATFEASDLERAQKVALNFTSSNQIGLTKTDVTDYHIDVSNLVACLSDFECLDE 671

QY 696 KRELLEKVKYAKRLSDERNLQDNFTSINKQPDFISTNEQSNFTSIHQSHGWSGSN 755

Db 672 KRELSEKVKHAKRLSDERNLQDNFNGRGINQPD-----RGMRGSTD 713

QY 756 ITIQEGNDVFKENYVTLPGTFNECYPTLYYOKIGESSELKAYTRYOLRGVIEDSODLEIYL 815

Db 714 ITIQGGDDVFKENYVTLPGTFVDECYPTLYYOKIDESKUKAYTRYELRGVIEDSODLEIYL 773

QY 816 IRYNAKHETLDPVGTESVWPLSVESPIGRCEPNRCAPHFENWPNPLDCSCRDGECARHS 875

Db 774 IRYNAKHETLDPVGTESVWPLSVESPIGRCEPNRCAPHFENWPNPLDCSCRDGECARHS 833

QY 876 HHFSLDIDIGCTDLHENLGVWVVFVKIKTQEGHARLGNLEFIEBEKPLLGEALSVRKRAEK 935

Db 834 HHFTLIDIVGCTDLNEDLGVWVVFVKIKTQEGHARLGNLEFIEBEKPLLGEALSVRKRAEK 893

QY 936 WRDKREKLOLETXYTEAKEAVDALFVDSQYNRLQADTNGIMHAAKLVHRIREAYLS 995

Db 894 WRDKREKLOLETXYTEAKEAVDALFVDSQYNRLQADTNGIMHAAKLVHRIREAYLS 953

QY 996 ELSVTPGVNARIPELEGRIITATSLYDARNVKNNGFNGLACVNVKHGVDV-QQSHHR 1054

Db 954 ELSVTPGVNARIPELEGRIITATSLYDARNVKNNGFNGLACVNVKHGVDV-QQSHHR 1013

QY 1055 SVLVIPEWAEVSVQAVRVCPCRGYILRTAYKKEGEGCVTIHEIENNTDELKFNCEBE 1114

Db 1014 SVLVIPEWAEVSVQAVRVCPCRGYILRTAYKKEGEGCVTIHEIENNTDELKFNCEBE 1073

QY 1115 EYPTDTGTNDYTA-----HOGTAVCSNRNAGYEDAYEVDVTASVNYKPTTBEETTYDVR 1170

Db 1074 EYPTDTGTNDYTA-----HOGTAVCSNRNAGYEDAYEVDVTASVNYKPTTBEETTYDVR 1131

QY 1171 RDNCEHCEYRGVYVYPPPLPAGYVTKOLEYFPETDKVWIEIGETEGTIVDSVLELLMEE 1228

Db 1132 RENCESNRGYDTPPLPAGYVTKOLEYFPETDKVWIEIGETEGTIVDSVLELLMEE 1189

RESULT 13

US-09-972-175-61

Sequence 61, Application US/09972175

Publication No. US20030101482A1

GENERAL INFORMATION:

APPLICANT: Baum, James A.

Gilmer, Amy Jelen

Mettus, Anne-Marie Light

TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:


```

Db 393 --TPNSFTYGRGTV-----DSLTELPPEDNSVPPREGYSHRLCHATFVQ 436
Qy 478 QSRVHV-----PVYSWTHRSADRTNTISSDSITQIPLVKSFNLNGTSVWSGPGFTGGDI 533
Db 437 RSGTFFLTGTVFVSWTHRSATLNTIDPERINQIPLVKGFRVWGGSVITGPGFTGGDIL 496
Qy 534 RTNVGSLVSLMGLNFNNNTSLQRVVRVYAAO-----TMVLRVTVGSTTFDQGFPFS 586
Db 497 RRNTFGDFVSLQVNNINSPIQRYRLRFRYASSRDARVIVLTGAASTGVGGQVSVNNPLQK 556
Qy 587 TMSANESLTSQSRFAFF-----VGISASGSQTAGISISNNAGRQTFHFDKIE 635
Db 557 TMEIGENLTSRTFRYTDFSNPPSFRANPDIIIGISEQPLFGAG-SISSG-----ELYDKIE 611
Qy 636 PIPITATFEAYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDDEFCLDE 695
Db 612 IILADATFEABSLERAQAVNALFTSSNQIGLTKTDVTDYHIDQVSNLVCLSDDEFCLDE 671
Qy 696 KRELLEKVKYAKLSDERNLLQDPNFTSINKOPDFTSINEQSNFTSIHQSEHGWGSEN 755
Db 672 KRELSEKVKHAKLSDERNLLQDPNFRGINRQPD-----RGMRGSTD 713
Qy 756 IITQEGNDVFKENYVTLPGTFNECYPTLYQKIGSELKAYTRYQLRGYIEDSODLEIYL 815
Db 714 IITQGGDDVFKENYVTLPGTVDECYPTLYQKIDESKLKAYTRYELRGYIEDSQDLEIYL 773
Qy 816 IRYNAKHETLDVPGTSEVPLSVESPIGRGCEPNRCAPHFENWPDLDSCSDRGKCAHHS 875
Db 774 IRYNAGHEIVNPGTGSMLPLSAQSPIGKCGEPNRCAPHLEWNPDLDCSDRGKCAHHS 833
Qy 876 HHFSLDIDIGCTDLHENLGVVVFVKIKTQEGHARLGNLEFIEEKPLLGEALSRVKRAEKK 935
Db 834 HHFTLIDIVGCTDLNEDLGVVVFVKIKTQDGHARLGNLEFLEEKPLLGEALARVKRAEKK 893
Qy 936 WRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRLQADTNIGMHAADKLVHRIREAYLS 995
Db 894 WRDKREKLQLEINIVYKAKESVDALFVNSQYDRLOQVDTNIAHAAADKRVHRIREAYLP 953
Qy 996 ELSVIPGVNAEIFELEGRIITAIISLYDARNVKNNGDFNGLACWNVKGVHVDV-QQSHHR 1054
Db 954 ELSVIPGVNAEIFELEGRIIFTAYSLYDARNVKNNGDFNGLLCCWNVKGVHVDVEEQNNHR 1013
Qy 1055 SVLVIPWEAEVSQAVRVCGRGYILRVYAYKEGYGEGCVTIHIEINTDELKPKNCEE 1114
Db 1014 SVLVIPWEAEVSQEVRCGRGYILRVYAYKEGYGEGCVTIHIEINTDELKFSNCEE 1073
Qy 1115 EYVPTDTCNDYTA---HOGTAVCNARNAGYEDAYEDVDTASVNYKPTYEEETVDYR 1170
Db 1074 EYVFNNTVTCNNYTGTOBEYEGT--YTSRQGYDEAYCNPNPSVPADYASVIEEKSYTDGR 1131
Qy 1171 RDNHCEYDRGVNYPPLPAGWMTXELBYFPETDKWIEIGETGKFIVDSVELLMEE 1228
Db 1132 RENPCESNRGYDGYTPLPAGVYTKDLEYFPETDKWIEIGETGTFIVDSVELLMEE 1189

```

Search completed: November 21, 2004, 17:04:22
Job time : 119 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 16:50:34 ; Search time 43 Seconds
(without alignments)
2747.772 Million cell updates/sec

Title: US-10-614-524-2
Perfect score: 6479
Sequence: 1 LTSNRKNEIINALSIPAV.....ICETGKEFIVDSVILLMEE 1228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5745	88.7	1228	2	S00873 parasporal crystal
2	3791	58.5	1176	2	A48970 parasporal crystal
3	3484.5	53.8	1189	2	S00944 parasporal crystal
4	3433.5	53.0	1165	2	S11446 parasporal crystal
5	3433	53.0	1166	2	S32645 parasporal crystal
6	3426.5	52.9	1181	2	A41052 parasporal crystal
7	3421	52.8	1171	2	I40572 parasporal crystal
8	3413.5	52.7	1174	2	S32649 parasporal crystal
9	3413.5	52.7	1176	2	JT0241 parasporal crystal
10	3412.5	52.7	1174	2	A42459 parasporal crystal
11	3411	52.6	1160	2	S32647 parasporal crystal
12	3405.5	52.6	1176	2	JC2219 parasporal crystal
13	3400.5	52.5	1176	2	S02215 parasporal crystal
14	3399.5	52.5	1176	2	A22617 parasporal crystal
15	3399.5	52.5	1178	1	USBSXH parasporal crystal
16	3399	52.5	1171	2	A37829 parasporal crystal
17	3387	52.3	1177	2	A49785 parasporal crystal
18	3323.5	51.3	1172	2	S32689 parasporal crystal
19	3283	50.7	1155	2	JD0002 parasporal crystal
20	3269	50.5	1155	2	A25513 parasporal crystal
21	3268	50.4	1155	2	S02134 parasporal crystal
22	3257.5	50.3	1156	2	A29125 parasporal crystal
23	3251	50.2	1155	2	I39838 parasporal crystal
24	3161.5	48.8	1156	2	A29838 parasporal crystal
25	2434	37.6	934	2	A22798 parasporal crystal
26	2294.5	35.4	1138	2	A48944 parasporal crystal
27	2200.5	34.0	1157	1	S49247 parasporal crystal
28	2170.5	33.5	719	2	I40590 cryv465 protein -
29	2116	32.7	719	2	I39815 insecticidal prote

RESULT 1

S00873

Parasporal crystal protein cryBa1 - Bacillus thuringiensis subsp. thuringiensis
N:Alternate names: parasporal crystal protein cryA4
C:Species: Bacillus thuringiensis subsp. thuringiensis
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S00873

R:Brizzard, B.L.; Whiteley, H.R.
Nucleic Acids Res. 16, 2723-2724, 1988
A:Title: Nucleotide sequence of an additional crystal protein gene cloned from Bacillus
A:Reference number: S00873; MUID:88203216; PMID:3362680
A:Accession: S00873
A:Molecule type: DNA

A:Residues: 1-1228 <BRI>
A:Cross-references: UNIPROT:P05517; EMBL:X06711; NID:g40264; PIDN:CAA29898.1; PID:g58094
C:Genetics:
A:Gene: cryA4
A:Start codon: TTG
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match	88.7%;	Score 5745;	DB 2;	Length 1228;
Best Local Similarity	89.1%;	Pred. No. 0;		
Matches 1098;	Conservative 35;	Mismatches 90;	Indels 10;	Gaps 3;
Qy	1	LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCTAEGNNINPLVSASTVQTGI	60	
Db	1	MTSNRKNEIIN-----AVSNHSAQMDLLDPDARIEDSLCTAEGNNIDPFVSASTVQTGI	55	
Qy	61	NIAGRIILGVLPAGQIASFVFLVGLWPRGRDQWEIFLEHVQLINQQITENARNTA	120	
Db	56	NIAGRIILGVLPAGQIASFVFLVGLWPRGRDQWEIFLEHVQLINQQITENARNTA	115	
Qy	121	LARLQGLGDSFRAYQQSLEDWLENRRDARTSVLYTQYIALELDFLNAMPLFAIRNQVEVP	180	
Db	116	LARLQGLGDSFRAYQQSLEDWLENRRDARTSVLYTQYIALELDFLNAMPLFAIRNQVEVP	175	
Qy	181	LLMVAQAANLHLLLRDASLFGSEFGLTSEIQRVYERQVEQTRDYSDYCVWNTGLN	240	
Db	176	LLMVAQAANLHLLLRDASLFGSEFGLTSEIQRVYERQVEQTRDYSDYCVWNTGLN	235	
Qy	241	SLRGNTAASWRYNQFRDLTLGLDLVALPSPVDTRTPINTSAQLTRVYTTDAIGATG	300	
Db	236	SLRGNTAASWRYNQFRDLTLGLDLVALPSPVDTRTPINTSAQLTRVYTTDAIGATG	295	
Qy	301	VNMASMNWYNNNAPSFAIAETAVIRSPHLLDFLEQLTIFSTSSRWASATRHMTYWRGHTIQ	360	
Db	296	VNMASMNWYNNNAPSFAIAEAAIRSPHLLDFLEQLTIFSASSRWASATRHMTYWRGHTIQ	355	
Qy	361	SRPIGGGLNTSTHGSGTNTSINPVRLSFPFRDVYWTESYAGVLLWGIYLEPIHGVTVPFN	420	
Db	356	SRPIGGGLNTSTHGATNTSINPVTLRFASRDVYRTESYAGVLLWGIYLEPIHGVTVPFN	415	

ALIGNMENTS

30	2107	32.5	719	2	S25383	parasporal crystal
31	2101	32.4	719	2	I39814	insecticidal prote
32	2002	30.9	1160	2	I40589	parasporal crystal
33	1947.5	30.1	1154	2	S39536	parasporal crystal
34	1930	29.8	823	2	S04181	parasporal crystal
35	1621	25.0	1156	2	S19306	parasporal crystal
36	1616.5	24.9	1136	1	USBS81	parasporal crystal
37	1511.5	23.3	1180	2	I39870	parasporal crystal
38	1498.5	23.1	1180	2	A26858	parasporal crystal
39	1227	18.9	380	2	B42459	hypothetical prote
40	1164	18.0	655	2	JC7140	protoxin - Bacilli
41	1064	16.4	652	2	A27323	parasporal crystal
42	1062.5	16.4	659	2	S10228	parasporal crystal
43	1053.5	16.3	652	2	I39811	parasporal crystal
44	999.5	15.4	649	1	JH0261	parasporal crystal
45	987.5	15.2	618	2	S11445	parasporal crystal

QY 969 RLOADTNIGMHAADKLVHRIIRAYLSLSVTPGVNAEIPFEELEGRIITATSLYDARNVV 1028
DB 914 RLQADTNIAIHAADKRVHRIIRAYLPELSVTPGVNAGIPFEELEGRIIFTAYSLYDARNV 973
QY 1029 KNGDFNGLACNNVKGHDV-QQSHRSVLVTPWEAEVSQAVRVCPCRGYILRTYAKE 1087
DB 974 KNGDFNGLACNNVKGHDVVEBQNNHRSVLVPEWEAEVSQAVRVCPCRGYILRTYAKE 1033
QY 1088 GXGEGCVTTHIENNTDELKFNCEEEVYPTDGTGNDYTAHQ---GTAVCSNRNAGYE 1144
DB 1034 GXGEGCVTTHIEDNTDELKFNCEEEVYPNVTCNDYATATQEBYGAY-TSRNAGVG 1092
QY 1145 DAYEVDTTASVNVKPYREBETTVDRDNHCEYDRGVNVPPLPAGYMTKELEYFPTDK 1204
DB 1093 KSYESNSVQADYASYEBKADTQGRDNHCHESNRGYGDTPLPAGYVTKLEYFPTDK 1152
QY 1205 VWIEIGETEGKFIIVSVLELLMEE 1228
DB 1153 VWIEIGETEGTIVDSVLELLMEE 1176

RESULT 3
S00944
parasporal crystal protein cry1Ca1 - Bacillus thuringiensis (strain entomocidus 60.5)
C:Species: Bacillus thuringiensis
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S00944
R:Honee, G.; van der Salm, T.; Visser, B.
Nucleic Acids Res. 16, 6240, 1988
A:Title: Nucleotide sequence of crystal protein gene isolated from B. thuringiensis subsp. israelensis
A:Reference number: S00944; MUID:88289380; PMID:3399402
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1189 <HON>
A:Cross-references: UNIPROT:P05518; EMBL:X07518; NID:g40293; PIDN:CAA30396.1; PID:g40294
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 53.8%; Score 3484.5; DB 2; Length 1189;
Best Local Similarity 56.8%; Pred. No. 1.4e-213; Indels 109; Gaps 20;
Matches 715; Conservative 143; Mismatches 291;

QY 7 NENEI--NALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGINIAG 64
DB 5 NQNCIPYNCLS-----NPBEVLDDGERISTGN-----SSIDISLSLVQ 43
QY 65 RILGVLGVPPAGQIASFYSFLVGLWPRGRDQWEIFLEHVEQLINQQTENARNATALARL 124
DB 44 FLVSNF-VPGGFLVGLIDFVWGIQV---SQWDAFLVQIEQLINERIAEFARNAAIANL 99
QY 125 QGLGDSFRAYQOSLELDLENRDDARTSRVLYTOYIALELDPLNAMPLEAIRNQEVPLLMV 184
DB 100 EGLGNFNFIYEAPEWEEDPNPPTRTVIDRIFRILGDLERDIPSPRISGFVPLLSV 159
QY 185 YAAANLHLLARLSDALFSGSEGLTSQEIQRYYERQVEQTRDYSQVWYNTGLNSLRG 244
DB 160 YAAANLHLLARLSDVIFGFRGLTTINVENYNNLRIRHIDVEADHCANTYNNRGLNLPK 219
QY 245 TWAASWRYNQFRDLTLGLVDLVALPSPYDTRTPINTSAQLTREVTDAIGATGVMA 304
DB 220 STYQDMITYNRLRDLTLVLDAIAFFPNYDNRYPYIQPVGOLTRVETDPL-----INFN 275
QY 305 SMNWNNAAPSATETAVIRPHLLDLEQLTIFSTSSWSATRHMTYWRGHTIQSRPI 364
DB 276 POLQSVQALPTFNWESSRIRNPHLDILNLTIFD---WPSVGRNFWGHRVVISLI 332
QY 365 GCGLNTSTHGSTNTINPVLRSFFGRDVYWTESYAGVLL-----WGIYLEPIHGVPTRFN 420
DB 333 GCGNITSPYIGREANQEPSPFTFNGPVPFTLSNLTLLRLLQQPWPAPPNLRGVGVFS 392
QY 421 FRNPONTFE---RGYANYSQPYESPLQLKDSLETLPPTTRPNYRESYSHRLSHIGLS 477

DB 393 --TPTNSFTYGRGTV-----DSLTELPPEDNSVPPREGYSHRLCHATFVQ 436
QY 478 OSRVHV---PVYSWTHRSADRTNTISDSSITQIPLVKSEFNLSNGTSVSVSGPFGTGDII 533
DB 437 RSGTFPLTGVVFSWTSASALTNTIDPERINQIPLVKGFRVWGTSVITGPGFTGDL 496
QY 534 RTNVNGSVLSMGLNFNNTSLQRYRVRVRYAASQ-----TMVLRTVTGGSTTFDQGFPS 586
DB 497 RRTTGDGDFVSLQVNLINSPTQRYLRFRVASSRDARVILTGAASGTGCGVGVSVNPLQK 556
QY 587 TMSANESLTSQSFRABFP-----VGISASGQTAGISISNNAGROTFFHPDKIE 635
DB 557 TMEIGENTSTFRYTFDFSNPFPFRANPDIIIGISEQPLFGAG-SISGG-----ELYDKIE 611
QY 636 FPIPTATFEAEVDLERAQEAVALNFTNTPRELKTDVTDYHIDQVSNLVACLSDEPCLDE 695
DB 612 ILLADATFEAESDLERAQKAVNALFTSSNQGLKTDVTDYHIDQVSNLVACLSDEPCLDE 671
QY 696 KRELLEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHESQSHGWSGN 755
DB 672 KRELSEKVKHAKRLSDERNLLQDPNFRGINRQPD-----RGMRGSTD 713
QY 756 ITIQENDVFKENYVTLPGTFNECYPTLYYQKIGESLKYATRYQIRGYIEDSQDLEIYL 815
DB 714 ITIQGGDDVFKENYVTLPGTFDECYPTLYYQKIDESLKYATRYELRGYIEDSQDLEIYL 773
QY 816 IRYNAKHETLDVPGTGVPLSVESPIGRCEPNRCAPHFENWPDLDSCRCRGEKCAHHS 875
DB 774 IRYNAKHIVNVPGTGSVPLVSAQSPICKCEPNRCAPHLEWPDLDSCRCRGEKCAHHS 833
QY 876 HHFSLDIDIGCTDLHENIGVMVVFVKIKTQSGHARLGNLEFIEBEKPLLGEALSrvKRAEKK 935
DB 834 HHFTLIDIVGCTDLNEDLGVMVFPKIKTQDGHARLGNLEFIEBEKPLLGEALSrvKRAEKK 893
QY 936 WRDREKQLQETKRVYTTAKAVDALFVDSQVNRLOADNTNIGMHAADKLVHRIIRAYLS 995
DB 894 WRDREKQLQETNIVYKEAKESVDALFVNSQYDLQVDTNIAIMIHAADKRVHRIIRAYLS 953
QY 996 ELSVTPGVNAEIPFEELEGRIITATSLYDARNVKNNGDNGLACNNVKGHDV-QQSHHR 1054
DB 954 ELSVTPGVNAEIPFEELEGRIIFTAYSLYDARNVKNNGDNGLACNNVKGHDVVEBQNNHR 1013
QY 1055 SVLVIPEAEVSQAVRVCPCRGYILRTYAKYEGGECVTHIENNTDELKFNCEEE 1114
DB 1014 SVLVIPEAEVSQAVRVCPCRGYILRTYAKYEGGECVTHIENNTDELKFNCEEE 1073
QY 1115 EYPTDGTGNDYTA-----HGTAVCSNRNAGYEDAYEDVTASVNYKPYREBETTVDR 1170
DB 1074 EYPPNNTVTCNNYTGTOEEYEGT--YTSRNGQYDEAYGNPNFSPADYASVYBEKSYTDGR 1131
QY 1171 RDNHCEYDRGVNVPPLPAGYMTKELEYFPTDKVWIEIGETEGKFIIVDSVLELLMEE 1228
DB 1132 RENPCESNRGYGDTYPLPAGYVTKDLEYFPTDKVWIEIGETEGKFIIVDSVLELLMEE 1189

RESULT 4

S11446
parasporal crystal protein cry1Da1 - Bacillus thuringiensis
N:Alternate names: Parasporal crystal protein cry1D
C:Species: Bacillus thuringiensis
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S11446
R:Hoefte, H.; Soetaert, P.; Jansens, S.; Peferoen, M.
Nucleic Acids Res. 18, 5545, 1990
A:Title: Nucleotide sequence and deduced amino acid sequence of a new Lepidoptera-specific
A:Reference number: S11446; MUID:91016842; PMID:2216728
A:Accession: S11446
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1165 <HOE>
A:Cross-references: UNIPROT:P19415; EMBL:X54160; NID:g40279; PIDN:CAA38099.1; PID:g40280
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 53.0%; Score 3433.5; DB 2; Length 1165;
Best Local Similarity 57.4%; Pred. No. 2.4e-210;
Matches 703; Conservative 141; Mismatches 273; Indels 107; Gaps 22;

QY 40 CIARGNNINPLVSASTVQTGINTAGRTGLVGLG-----VPPAQIASFSPFLVGLWPRG 93
DB 14 CLSNPKEI--ILGERLETGNTVADISGLINFLYFNFPVGGGFVGLLELTWGTGP-- 69

QY 94 RDQWEIEFLEHVQIINQITENARNTALARLOGLGDSFRAYQOQSLDMLNDRDDARTSV 153
DB 70 -SQWDIFLAQIEQLISORIEBEPARNOAIGRLGSLNLYKVVYVAFSDWEKDTNPALREE 128

QY 154 LYQYTALEDDFNAMPLFAIRNOEYPLLWYQAANLHLLLRDASLFGSEFGTSQEI 213
DB 129 MRLOFNDMSALTALPLFRQVNEYVALLSVYQAANLHLSILRDSVFGERWGYDTATI 188

QY 214 QRYERQVQTRDYSYCVENYNTGLNSLRGTNAASWRYNQFRDLTLGLVDLVALPPS 273
DB 189 NNRYSDLTSLIHVYTHNCVDYTNQGLRLLEGRFLSDWYVNRFRQLIISVLVDIVAFPPN 248

QY 274 YDRTYTPINTSAQLTREVVYTDGATGATGVMAWNNWYNNNA-----PSFSAIETAVIRSP 327
DB 249 YDRTYTPIQATQLTRVYLD-----LPPINENLSPAASYPTFSAESAIIIRSP 297

QY 328 HLLDFLEQLTIFSTS--SRWSATRHMTYWRGHTIQSRPIGGGLN---TSTHGSTWTSINPV 383
DB 298 HLVDLFLNSFTIIVTDSLARYA-----YGGHLVNSFRGTNTNLRSLIRSLYREGNTERPV 351

QY 384 RLSFF--SRDYYTESVAGVLLWGI--YLEPIHGVPVTRFNFRNPQNTFERGTANYSQPVES 441
DB 352 TITASPSVPDIFRTLSY-----ITGLDNSNPVAGIEGVF-----QNTISRSYRKS GPI-- 400

QY 442 PGLQLKDSLETLPPTTERPNYESYSHRLSH---IGLISQSRVHVVPYVSWTHRSADTNT 498
DB 401 -----DSPSELPPQDASVSPAIGYSHRLCHATFLERISGPRIAGTVFSWTHRSASPTNE 454

QY 499 ISSDSITQIPLVKSFLNLSGTSVSGPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQYRV 558
DB 455 VSPSRITQIPWYKAHTLASGASVINGPGFTGGDILTRNSMGLTGLVTTGRLPQSYII 514

QY 559 RYRYA--ASQTVLWLVTVGGSTFDQGPSTWMSANESITQSFRFAEPFVGIASGSG-- 614
DB 515 RFRYASVANRSGTRYSQPPS--YGISFPKTMDSAGEPLTGRSFAHTLFTPTIFSAQEE 572

QY 615 -----TAGISNNAGRTFHDKIEFIPITATPEASYDLERAQEAVALFTNTNPRLK 669
DB 573 FDLYIQSGV-----YIDRIEPIPTATPEASYDLERAQKVVALFTNTNQLGLK 621

QY 670 TDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQDPNFTSINKQPD 729
DB 622 TDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQDPNFRGINRQPD 681

QY 730 FISTNEQSNTSIHQSEHGMWGSSENTIQEGNDVFKENVVTLPGTFNECYPTLYQKIG 789
DB 682 -----RGMWGSTDITIQGGDDVFKENVVTLPGTFDECIPTLYQKID 723

QY 790 ESELKAYTRYOLRGYIEDSDQLEIYLRNAKHETLDVPGTESVWPLSVESPIGRGEPN 849
DB 724 ESKUKATRYOLRGYIEDSDQLEIYLRNAKHEIVNVPGTGLMPLSVENQIGPCGEPN 783

QY 850 RCAPHFENFPDLCSDRGDKCAHSHHFLSDIDIGCTDLHENLGVVWVFKITQEGHAR 909
DB 784 RCAPHLEWNPDLHCSDRGDKCAHSHHFLSDIDVGTDLNEDGLGVWVIFKIKTQGHAR 843

QY 910 LGNLEFTEEPPLAGEALSVRKRAEKWRDKREKQLQETKRVYTEAKAEVADLFVDSQVNR 969
DB 844 LGNLEFTEEPPLAGEALSVRKRAEKWRDKRETLQLETTIVYKEAKESVDALFVNSQYDR 903

QY 970 LQADTNIGMTHAADKLVHRIREAYLSLSVPGWNAEIEFEELGRIITAIISLYDARNVK 1029
DB 904 LQADTNIAMHAAKRVHRIREAYLDELVSIVPGVNAIIEFEELRIFTAFSLYDARNIIK 963

QY 1030 NGDFNNGLACWNVKGVHDV--QOHSRSLVLIPEWEAEVSAQVRCVPGRGYILRVATYKEG 1088
DB 964 NGDFNNGLLCWNVKGHVEVEBQNNHRSVLVIPEWEAEVSAQVRCVPGRGYILRVATYKEG 1023

QY 1089 YGECVTHIHEIENNTDELKFNKCEEEVYPTDGTGNDYTA---HGTAVACNRSNAGYE 1144
DB 1024 YGECVTHIHEIENNTDELKFNKCEEEVYPTDGTGNDYTA---HGTAVACNRSNAGYE 1081

QY 1145 DAYEVDVTASVNYKPYTEESTYTDVDRDNHCEYDRGVYVPPPLPAGYMTKELYFFETDK 1204
DB 1082 EAYGNNSVPADVASVYEKSYTDRRENPCSNRGVGYTPLPAGVTKLELYFFETDK 1141

QY 1205 VWIEIGETGKFIYVDSVELLMEE 1228
DB 1142 VWIEIGETGKFIYVDSVELLMEE 1165

RESULT 5
S32645
parasporal crystal protein cryiGal - Bacillus thuringiensis
C/Species: Bacillus thuringiensis
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C/Accession: S32645
R/Lambert, B.
submitted to the EMBL Data Library, April 1993
A/Reference number: S32645
A/Accession: S32645
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1166 <LMA>
A/Cross-references: UNIPROT:Q45746; EMBL:222510; NID:g295861; PIDN:CAA0233.1; PID:g29586
C/Superfamily: parasporal crystal protein
C/Keywords: delta-endotoxin

Query Match 53.0%; Score 3433; DB 2; Length 1166;
Best Local Similarity 57.9%; Pred. No. 2.5e-210;
Matches 693; Conservative 137; Mismatches 281; Indels 86; Gaps 19;

QY 54 STVQTGINAGILGVLPFAGQIASFYSLVGLMWRGRDQWEIFLEHVQIINQIT 113
DB 34 SQVSSGLT---RFLLEAAVPEAGFALGLFDIIGWAL---GVDQNSLFLRQIEQLRQEI 87

QY 114 ENARNTALARLQGLGDSFRAYQOQSLDMLNDRDDARTSVLYTQYIALELDFLNAFLPA 173
DB 88 ELERNRATAILTGLSSSLNLYVEALREWENDPNPNPASQERVTRFRLTDDAIVTGLEPTLA 147

QY 174 IRNQEVPLLMYQAANLHLLLRDASLFGSEFGITSOEIQRYVERQVQTRDYSVCVE 233
DB 148 IRNLEVNLSVYTOAANLHLLLRDASVYFGRWGLTQANIEDLYTRLTNSIQEYSDHCAR 207

QY 234 WYNTGLNSLRGTNAASWRYNQFRDLTLGLVDLVALPSPYDTRTYPTINTSAQLTREYVT 293
DB 208 WYNTGLNSLRGTNAASWRYNQFRDLTLGLVDLVALPSPYDTRTYPTINTSAQLTREYVT 263

QY 294 DAIGATGVNMAWNNWYNNAPSPSAIETAVIRSPHLLDLEQLTIFSTSRWSATRHMTY 353
DB 264 SPVWAGIN-----FGLSTANVLRAHLMDFIDRVIYTN-----VRETPY 305

QY 354 WRGHTIQSRPIGGG-----LMTSTHGSTWTSINPVL---SFFSRDYYTESVAGVLLWGI 406
DB 306 WAGHEVTSRRTGQGGNEIRFPYGVAAAEPPVTIRPTGTGTDEQRQWRARSRVVSFRS 365

QY 407 YLEPIHGVPVTRFNFRNPQNTFERGTANYSQPVESPGQLKDSLETLPPTTERPNYESY 466
DB 366 SQDQFSLVDVAGF-----LTIFSAVSIYANGF---GFN-TDTIDEIPIEGTD--PFTGY 413

QY 467 SHRLSHLGLISQS-----RVHVVPYVSWTHRSADTNTISGDSITQIPLVKSFLNLSGTSV 521
DB 414 SHRLCHGVFLASGPFISQYARAPIFSWSHRSATLTNTIAPDVITQIPLVKAFNLHSGATI 473

QY 522 VSGPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQYRVVRYVAASQTMVLRVTVGGSTTF 580
DB 474 VKGPGFTGGDIILARTNV-GSFGDMRVNITAPLSQRVIRVIRVASTTDLQFYNTINGTTIN 532

Db 681 NLLQDPNPRGINRQLD-----RGMRGSDTITIQGGDDVFKENYVTL 722
Qy 774 GFNECYPTLYQKIGESLKYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDVPGTESV 833
Db 723 GTFDECYPTLYQKIDESLKYTRYQLRGYIEDSQDLEIYLIRYNAKHETVNPVGTSL 782
Qy 834 WPLSVESPIGRGEPNRCAPHPEWNPDLDCSCRDGKCAHSHHPSLDIDIGCTDLHENL 893
Db 783 WPLSAQSPFGKCGEPNRCAPHLEWNPDLDCSCRDGKCAHSHHPSLDIDIGCTDLNEDL 842
Qy 894 GVVVFKITQDCHARLGNLEFIEBKPLLGALSRVKAERKWRDKREKLOLETKRVVTE 953
Db 843 GVVVFKITQDCHARLGNLEFIEBKPLVGALSRVKAERKWRDKREKLOLETKRVVTE 902
Qy 954 AKEAVDALFVDSQYNRLQADTNIGMHAADKLVRHREAYLSLSVPGVNAEIEFELEG 1013
Db 903 AKESVDALFVDSQYNRLQADTNIGMHAADKLVRHREAYLSLSVPGVNAEIEFELEG 962
Qy 1014 RIITAIISLYDARNVVKNGDFNGLACMNVKGVHDV-QQSHRSVLVPEAEVSQAVRV 1072
Db 963 RIITAFSLYDARNVVKNGDFNGLSCMNVKGVHDVEEQNORSVLVPEAEVSQAVRV 1022
Qy 1073 CPGRGYILRVYAYKEGYGEGCVTIEIENNTDELKFNCEBEEVYPTDTGTCNDYTAHQ- 1131
Db 1023 CPGRGYILRVYAYKEGYGEGCVTIEIENNTDELKFNCEBEEVYPTDTGTCNDYTNQE 1082
Qy 1132 --GTAVCNRNAGVEDAYEVDTTASVNYKPYEBEETVDRDNHCEYDRGVNYPPLPA 1189
Db 1083 EYGGAY-TSRNRGYNAPSV----PADYASVYSEKSYTDGRENPCBFNRGYRDTPLPV 1137
Qy 1190 GYMTKELEFPETDKVWIEIGETGKFIIVDSVELLMEE 1228
Db 1138 GYMTKELEFPETDKVWIEIGETGTFIVDSVELLMEE 1176

RESULT 10
A42459
Paraaporal crystal protein cry1Fa1 - Bacillus thuringiensis (strain aizawai)
N:Alternate names: paraaporal crystal protein cry1F
C:Species: Bacillus thuringiensis
C>Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 09-Jul-2004
C:Accession: A42459
R:Chambers, J.A.; Jelen, A.; Gilbert, M.P.; Jany, C.S.; Johnson, T.B.; Gawron-Burke, C.
J. Bacteriol. 173, 3966-3976, 1991
A:Title: Isolation and characterization of a novel insecticidal crystal protein gene from
A:Reference number: A42459; MUID:91286178; PMID:2061280
A:Accession: A42459
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1174 <CHA>
A:Cross-references: UNIPROT:Q03746; GB:M63897; NID:g142757; PIDN:AAA22348.1; PID:g142758
C:Superfamily: paraaporal crystal protein
C:Keywords: delta-endotoxin

Query Match 52.7%; Score 3412.5; DB 2; Length 1174;
Best Local Similarity 58.4%; Pred. No. 5.2e-209;
Matches 693; Conservative 132; Mismatches 271; Indels 91; Gaps 21;

Qy 72 VPPAGQIASFVSFLVGLMPPRGDQWEIFLEHVEQLNQIQTENANTALRQGLGDSF 131
Db 49 VPGVGVAFLGFDLWIGFITP---SDSLFLQLEQLIEQRIETLERNRAITTLRGLADSY 105
Qy 132 RAYQOQSLDLEWLRDADRSLVLYTOYIALELDFLAMPFLAIRNOEVLPLMVAQAANL 191
Db 106 EIVTEALREWEANPNNAQLREDVRIRFANTDDALITAINFTLTSFEIPLLSVYQAANL 165
Qy 192 HLLILRDASLFGSEFGLTSQEIQRHYRQVEQTRDVSQYCVWEYNTGLNSLRGTNAASV 251
Db 166 HLLILRDASVFGQGWGLDIAIVANNHVRNLINLTHRYTKHCLDTYNOGLENLRGTNRQWA 225
Qy 252 RYNOFRDLTLGLVLDLVALPSPDYDTRYPTINTSAQLTREYVYTDATGATGNWASMNWYN 311
Db 226 RFNOFRDLTLVLDLVALPSPDYDTRYPTINTSAQLTREYVYTDATGATGNWASMNWYN 278

RESULT 11
S32647

paraaporal crystal protein cry1Db1 - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S32647
R:Lambert, B.

Qy 312 NAPS-FSAIETAVIRPHLLDFLEQLTIFSTSSRWASATRHMTYVRGHTIQSRPIGGG-LN 369
Db 279 NIPNGFNRAEFGV-RPHLMDFNLSL--FVTA---ETVRSQTWGGHLVSRMTAGNRIN 332
Qy 370 TSTHGSTNTSINPRLSPFS-----RDVYWTESVAGVLLMGYILEPIHGVTVPFRFNPNQ 425
Db 333 FPSYG-----VFNPGGAIWIADEDRPPEYRTLS-----DPVF-----VRGGFGNPH 373
Qy 426 N-----TPEGTANYSQPYESPGQLKDXSETLPEPETERPNYSYSYRHLGHI----- 473
Db 374 YVLGLRGVAFQQTGTHTRFRNSG--TIDSLDEIPQDNGSGAPMNDYSHVLNHTVFRW 431
Qy 474 -GLISOS-RVHPVYVSWTHRSADRTNTISDSITQIPLVKSFNLSNGTSVVSQPGTGGD 531
Db 432 PGBISGSDSWRAPMFWSWTHRSATPTNTIDPERITQIPLVKAHTLQSGTTVVRGPGTGGD 491
Qy 532 IIRTNVNGSVLSMGLNPNNTSLQRYRVRYAASQTMVLRTVGGSTTTTDDQGPSTMSAN 591
Db 492 ILRRTSGGPPAYTIVNINGQLPQRYRARIYASTTNLRIYVTVVAGERIFAGQFNKMTDTG 551
Qy 592 ESLSQSFRAPEPPVGISASGSQTAGISISNNAGROT-----HFDKIEFIPITAFEA 645
Db 552 DPLTFQSFVATINTAFTFMSQS-----SFTVGADTFSSGNEVYIDREFLIPVTAIFEA 606
Qy 646 EYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLKVKY 705
Db 607 EYDLERAQAVNALFTSIQIGIKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLKVKY 666
Qy 706 AKELSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQSEHGWGWSNITIOEGNDVF 765
Db 667 AKELSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQSEHGWGWSNITIOEGNDVF 708
Qy 766 KENYVTLPGTFNECYPTLYQKIGESLKYTRYQLRGYIEDSQDLEIYLIRYNAKHETL 825
Db 709 KENYVTLPGTFNECYPTLYQKIGESLKYTRYQLRGYIEDSQDLEIYLIRYNAKHETV 768
Qy 826 DYPGTSWPLSVESPIGRGEPNRCAPHPEWNPDLDCSCRDGKCAHSHHPSLDIDIG 885
Db 769 NVLGTSLMPLSVQSPFIRKCGEPNRCAPHLEWNPDLDCSCRDGKCAHSHHPSLDIDIG 828
Qy 886 CTDLHENLGVVVFVKITQDCHARLGNLEFIEBKPLLGALSRVKAERKWRDKREKLOL 945
Db 829 CTDLNEDLDVWIFKIKTDGCHARLGNLEFIEBKPLVGALSRVKAERKWRDKREKLEL 888
Qy 946 ETKRVYTEAKEAVDALFVDSQYNRLQADTNIGMHAADKLVRHREAYLSLSVPGVNA 1005
Db 889 ETNIVYKEAKESVDALFVNSQYDQLQADTNIGMHAADKLVRHREAYLSLSVPGVNV 948
Qy 1006 EIFELEGRITTAISLYDARNVVKNGDFNGLACMNVKGVHDV-QQSHRSVLVPEWEA 1064
Db 949 DIFEELKGRIFTAFPLYDARNVVKNGDFNGLSCMNVKGVHDVEEQNORSVLVPEWEA 1008
Qy 1065 EYSQAVRVCPGRGYILRVYAYKEGYGEGCVTIEIENNTDELKFNCEBEEVYPTDTGTC 1124
Db 1009 EYSQAVRVCPGRGYILRVYAYKEGYGEGCVTIEIENNTDELKFNCEBEEVYPTDTGTC 1068
Qy 1125 NDYTAHQ---GTAVCNRNAGVEDAYEVDTTASVNYKPYEBEETVDRDNHCEYDRGY 1181
Db 1069 NDYTAHQEYGGAY-TSRNRGYNAPSV----PADYASVYSEKSYTDGRENPCBFNRGY 1127
Qy 1182 VNYPLPAGYMTKELEFPETDKVWIEIGETGKFIIVDSVELLMEE 1228
Db 1128 GYTPPLPAGYMTKELEFPETDKVWIEIGETGTFIVDSVELLMEE 1174

QY 954 AKEAVDALFVDSQYNRLQADTNIGMHAADKLVHRIEAYISELSVPGVNAEIPFEELEG 1013
D 1013
D 903 AKESVDALFVNSQYDQLQADTNIAHAAKXVHSIREAYLPELSVPGVNAEIPFEELEG 962
QY 1014 RIITAFSLYDARNVVKNGDFNGLACWNVKGHDV-QQSHHRSVLVPEWEAEVSQAVRV 1072
D 963 RIITAFSLYDARNVVKNGDFNGLACWNVKGHDVVEEQNQRSVLVPEWEAEVSQAVRV 1022
QY 1073 CPGRGYLLRVAYKEGEGCVTHIEINNTDELAKFCEREEVPTDTGTCNDVTAHQ- 1131
D 1073 CPGRGYLLRVAYKEGEGCVTHIEINNTDELAKFCEREEVPTDTGTCNDVTAHQ- 1131
D 1023 CPGRGYLLRVAYKEGEGCVTHIEINNTDELAKFCEREEVPTDTGTCNDVTAHQ- 1082
QY 1132 --GTAVCNRSNAGYDAYEVDVTTASVNYKPYTEEBEYTDVDRDNHCEYDRGVNYPPLPA 1189
D 1083 EYGGAY-TSRNRYNEAPSV---PADYASVVEEKSVDGRRNCPFNRGVRYDTPLPV 1137
QY 1190 GYMTKELEYFPETDKWIEIGTEGTFIVDSVELLMEE 1228
D 1138 GYVTKLEYFPETDKWIEIGTEGTFIVDSVELLMEE 1176

RESULT 14
A22617
paraaporal crystal protein cryIa1 - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C:Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 09-Jul-2004
C:Accession: A22617
R:Schnept, H.E.; Wong, H.C.; Whiteley, H.R.
J. Biol. Chem. 260, 6264-6272, 1985
A:Title: The amino acid sequence of a crystal protein from Bacillus thuringiensis deduced from the complementary DNA sequence
A:Reference number: A22617; NID:85207613; PMID:2581950
A:Accession: A22617
A:Molecule type: DNA
A:Residues: 1-1176 <SCH>
A:Cross-references: UNIPROT:P02965; GB:M11250; NID:g142764; PIDN:AAA22353.1; PID:g142765
C:Superfamily: paraaporal crystal protein
C:Keywords: delta-endotoxin

Query Match 52.5%; Score 3399.5; DB 2; Length 1176;
Best Local Similarity 56.2%; Pred. No. 3.5e-208;
Matches 696; Conservative 149; Mismatches 294; Indels 100; Gaps 24;

QY 27 MDLSPDARIEDSLCIAEGNNINP---LVSASTVOTG---INIAGRILGLV---GVPPFAGQ 77
D 1 MDNPNPNI---NECIPYCLNPNPEVGLGRIETGYTPDISLSLTOFLSEFPFGAG- 55
QY 78 IASVYSFVLVBWP-RGRDQWEIFLEHVEQLINQITENARTALRQLGLGDSFRAYQQ 136
D 56 ---FVLGLVDIIMGIFGSPQWDAPVQIEQLINQRIEFPARNOAISRLLEGLSNLYQIYAE 112
QY 137 SLEDWLENRDDARTSVLYTQYIALELDFANMPLFAIRNOEVLPMYAAQANLHLL 196
D 113 SFRWEADPTNPALKEEMRIQFNDMNSALTATPILAVQNYQVPLLSVYVQAANLHLSVL 172
QY 197 RDALFSGSEGLTQOEIQRYYERQVEQTRDYSYCVIEWNTGLNSLRGTNAASVRYNQF 256
D 173 RDVSVFGQWGFDAATINSRNDLTRICNTDVAIRVNTGLERWVGPDSDRWRYNQF 232
QY 257 RRDITLGLVDLVALPSPYDTRYPINTSAQLTREYITDAIGATGVNMAWNNYNNAPSF 316
D 233 RRELTLVLDLVALPSPYDTRYPINTSAQLTREYITDAIGATGVNMAWNNYNNAPSF 280
QY 317 SAIEAV---TRSHLLDPLQLIFTSRWSATRWYRGTIQSRPIGCG---LNT 370
D 281 RGMQAQRIEONRPHMLILNSITIT---DVHRGFNYSWGHQITASPVSFGSPFAP 335
QY 371 STHGSTNTSINPRLSPFSRDVYTES---YAGVLLW-GIVLEPHGVPTVRFNFRN--- 423
D 336 PLFGNAGNAAPVLVSLTGLIGFRLTSLPLRILGSPNNQELFVLGDGTFEFSALTT 395
QY 424 --PQNTF-ERGATANYSQPYESFGLQKDSLETLPPTTERPNYRSYSHRLSHGLISQS- 479
D 424 --PQNTF-ERGATANYSQPYESFGLQKDSLETLPPTTERPNYRSYSHRLSHGLISQS- 479

Db 396 NLPSTIYRQRGTV-----DSLVDIPQDNSVPPRACGFSHRLSHVMTLSQA 441
QY 480 ---RVHVPVSWTHRSADRNTWISSDSITQIPLVKSFNLSAGTSVSGPGTGGDIIRT 535
D 442 GAVTTLRAPTFWSQHRSAEFNNIIPSSQITQIPLTKSTNLGSGTSVVRKPGTGGDILRR 501
QY 536 NVNGSVLSMGLFNNTSLQRVVRVRYAASQTMVLRVVGGSTTFDQGFPPSTMSANESLT 595
D 502 TSPGQISTLRVNIATPLSORVVRIRVASTTNLQPHSTSIDGRPINQGNFSAATSSGSLN 561
QY 596 SQSRFPAPFPVGIS-ASGQTAGISIS-NNAGROTFFHFKIEFIPITATFEAYDLERAQ 653
D 562 SGSPRTTGTFTFPFNGSGSVFTLSAHVFNSENVY-IDRIEFVPAEVTFEAYDLERAQ 620
QY 654 EAVNALFTNTPRRLKTDVTDYHIDQVSNLACLSDEFCLDEKRELLKVKYAKLSDBR 713
D 621 KAVNELFTSSNQIGLTKTDVTDYHIDQVSNLVECLSEDEFCLDEKRELLKVKYAKLSDBR 680
QY 714 NLLQDPNFTSINKQPDFISTNEQSNFTSIHQSEHGWMGSENITIQEGNDVFKENYVTL 773
D 681 NLLQDPNFRGINRQLD-----RGWRGSTDTITIQGGDDVFKENYVTL 722
QY 774 GTFNECYPTLYYQKIGSELSKAYTRYQLRGYIEDSQDLEIYLIRYNAXHETLDVPGTSV 833
D 723 GTFDECYPTLYYQKIDESKAYTRYQLRGYIEDSQDLEIYLIRYNAXHETVNVPGTSL 782
QY 834 WPLSVESPIGRCGPNRCAPHFNPNLDLDCSCROGEKCAHSHHPSLDDIDGCTDLHENL 893
D 783 WPLSAQSPIGKCGBNRCAPHLEWNPDLDCSCROGEKCAHSHHPSLDDIDGCTDLNEDL 842
QY 894 GWVWFKIKTQEGHARLGNLFIEBEKPLGLGALSRVKAERKWRDKRKLQLETKRVYTE 953
D 843 GWVWFKIKTQEGHARLGNLFIEBEKPLGLGALSRVKAERKWRDKRKLQLETKRVYTE 902
QY 954 AKEAVDALFVDSQYNRLQADTNIGMHAADKLVHRIEAYISELSVPGVNAEIPFEELEG 1013
D 903 AKESVDALFVNSQYDQLQADTNIAHAAKXVHSIREAYLPELSVPGVNAEIPFEELEG 962
QY 1014 RIITAFSLYDARNVVKNGDFNGLACWNVKGHDV-QQSHHRSVLVPEWEAEVSQAVRV 1072
D 963 RIITAFSLYDARNVVKNGDFNGLACWNVKGHDVVEEQNQRSVLVPEWEAEVSQAVRV 1022
QY 1073 CPGRGYLLRVAYKEGEGCVTHIEINNTDELAKFCEREEVPTDTGTCNDVTAHQ- 1131
D 1073 CPGRGYLLRVAYKEGEGCVTHIEINNTDELAKFCEREEVPTDTGTCNDVTAHQ- 1082
QY 1132 --GTAVCNRSNAGYDAYEVDVTTASVNYKPYTEEBEYTDVDRDNHCEYDRGVNYPPLPA 1189
D 1083 EYGGAY-TSRNRYNEAPSV---PADYASVVEEKSVDGRRNCPFNRGVRYDTPLPV 1137
QY 1190 GYMTKELEYFPETDKWIEIGTEGTFIVDSVELLMEE 1228
D 1138 GYVTKLEYFPETDKWIEIGTEGTFIVDSVELLMEE 1176

RESULT 15

USBSXH

paraaporal crystal protein cryIa1 [validated] - Bacillus thuringiensis subsp. kurstaki
C:Species: Bacillus thuringiensis subsp. kurstaki
C:Date: 18-Apr-1984 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A23962; A03489
R:Adang, M.J.; Staver, M.J.; Rocheleau, T.A.; Leighton, J.; Barker, R.F.; Thompson, D.V.
Gene 36, 289-300, 1985
A:Title: Characterized full-length and truncated plasmid clones of the crystal protein o
A:Reference number: A91526; MUID:86083171; PMID:3000881
A:Accession: A23962
A:Molecule type: DNA
A:Residues: 1-1178 <ADA>
A:Cross-references: UNIPROT:P05068; GB:M11068; NID:g142721; PIDN:AAA22331.1; PID:g142722
A:Experimental source: strain HD-73
A>Note: the authors translated the codon ATT for residue 11 as Leu
R:Wong, H.C.; Schnept, H.E.; Whiteley, H.R.
J. Biol. Chem. 258, 1960-1967, 1983

A;Title: Transcriptional and translational start sites for the Bacillus thuringiensis ch
A;Reference number: A92410; MUID:83109004; PMID:6296116

A;Accession: A03489

A;Molecule type: DNA

A;Residues: 1-76, 'P', 78-147, 'L', 149-247, 'S', 249-282, 'M', 284-285, 'R', 287-288, 'QN', 291-292

A;Experimental source: strain HD-1

C;Comment: This protein is present in crystalline form as a component of the spore coat.

C;Superfamily: parasporal crystal protein

C;Keywords: delta-endotoxin

```
Query Match      52.5%; Score 3399.5; DB 1; Length 1178;
Best Local Similarity 56.7%; Pred. No. 3.5e-208;
Matches 704; Conservative 148; Mismatches 287; Indels 102; Gaps 26;

QY 27 MDLSPDARTEDSLCAEGNNINP---LVSASTVQTG---INTAGRIILGV---GVPPAGQ 77
DB 1 MDNPNFNI---NECIPYNCLSNPEVEVLGGERIETGYTPIDISLSTQLTFLSEFVPGAG- 55

QY 78 IASFYSFLVGLWP- RGRDQWBIFFLEHVEQLINQOITENARTALARIQGLGDSFRAQQ 136
DB 56 ---FVLGLVDIIIGIFGPSQWDAFLVQIEQLINQRIEFARNQALISRLGLESLNYQIYAE 112

QY 137 SLEDWLENRDARTSVLTYQVIALELPLNMPLEFAIRNOEVPPLMYAQAANLHLLL 196
DB 113 SFREADPTNPALREEMRIQDNMNSALTTAIPLFAVQNTQVPLLSVYVQAANLHLSVL 172

QY 197 RDASLFGSEFGLTSQBIQRIYERQVEQTRDYSYCVWYNTGLNSLRTNAASVVRNQF 256
DB 173 RDVSFVGQRMGDAATINSRYNDLTRIGNTYDVAVRWYNTGLERVWGPDSDRVRYNQF 232

QY 257 RDRLTLGVLDIVALPFSYDTRTYPINTSAQLTREYITDAIGATGVNMASNMWYNNAPSF 316
DB 233 RRELTLTVLDIVALEPNDYRRPIRTVSQLTREIYT-----NPVLENFDCSPGSA 284

QY 317 SAETAVIRSPHLLDFLQLTIFSSRSWSATRWYTWRIGHTIOSRPIGG-----LNTSTH 373
DB 285 QCIERS-IRSPHLMILNITIIYT-----DAHGGYVWSHGQIMASPVFGSPGPEFTPLY 338

QY 374 GSTNTSINPVL-SFFSRDVTWTSYAGVLLWGLYLEPIHGVTVRNFRNPQ-----N 426
DB 339 GTMGNAAPQRIVAQLGGQVYRTLST-----LYRRPN-----IGINNQLSVLDGT 386

QY 427 TFERGTANY--SQVYESPGLQDKSETELPEPETERPNYESYSHRLSHIGLI-----S 477
DB 387 EFAGTSNLPSAVTRKSG--TVDSLDEIPEQNNVPPRQGFSHRLSHVSMFRSGFSNS 444

QY 478 QSRVHPVYVSWTHRADRTNTISSDSITQIPLVKSFNLSGTSVVGSGPFTGGDIIRTV 537
DB 445 VSIIRAPMFWIHSRAEFNNIIASDSITQIPAVKGNFLNG-SVISGPGFTGGDLVRLNS 503

QY 538 NGS-----VLSMGLNFNTSLQRYRVRYAASQTMVLRTVYGGSTTFDQFPFSTMSAN 591
DB 504 SGNINQNRGYIEVPIHFPTS-TRYRVRYASVTPPIHLNVNNGNSSIFNTVPATATSL 562

QY 592 ESLSQGFRAEFPVGIASGSPAGISINNAQRTFHFDEKIEFIPITATFAEYDLER 651
DB 563 DNLOSSDFGVPESANATSSLGNIVG--VRNFSGTAGVIIDREFIPVPTATLEAEYNLER 620

QY 652 AQEAVNALFTNTPRRLKTDVTYHIDQVSNLVACLSDDEFCLDEKLEKLVKAKRLSD 711
DB 621 AQKAVNALFTSTNQLGLKTNVTYHIDQVSNLVTYLSDEFCLDEKLEKLVKAKRLSD 680

QY 712 ERNLQDPNFTSNKQDPDFTSTNEQSNFTSIHQSEHGHWGSGENIITIQSGNDVFKENYVT 771
DB 681 ERNLQDSNFKDINRQ-----ERGMGGSTGITIQGGDDVFKENYVT 722

QY 772 LPGTFNECYPTLYQKIGSELKAYTRYQLRGYIEDSQLEIYLYRYNAKHETLDVPGTE 831
DB 723 LSGTFDECYPTLYQKIDSKLAFTRYQLRGYIEDSQLEIYLYRYNAGHETVNPVGTG 782

QY 832 SVMPLSVESPIGRCGPNRCAPHFENPDIDCSRDGKCAHHSHHFLSDIDIGCTDLHE 891
DB 783 SLWPLSAQSPIGKCGPNRCAPHLEWNPDLDCSRDGKCAHHSHHFLSDIDVGCTDLNE 842
```

Search completed: November 21, 2004, 17:02:15

Job time : 47 secs

```
QY 892 NLGVVVVFKIKTQSGHARLGNLSEIEKPLLGEALSRVKAEBKAKWRDKREKLQLETKRVY 951
DB 843 DLGVVWTFKIKTQSGHARLGNLSEIEKPLVGBALARVKAEBKAKWRDKREKLEWETNIVY 902

QY 952 TEAKEAVDALFVDSQYNRLQADTNIGMIHAADKLVHRIRAYLSLSVIFGVNAEIFEEL 1011
DB 903 KEAKESVDALFVNSQVDQLQADTNIAMHAADKRVHSIREAYLPESLVIQVNAEIFEEL 962

QY 1012 EGRITTAISLYDARNVVKNGDFNNGLACWNVKGHDV-QQSHHRSVLVIPWEAEVSAV 1070
DB 963 EGRIFTAFSLYDARNVVKNGDFNNGLSWNVKGHDVVEQNNQSVLVVPEWEAEVSEQV 1022

QY 1071 RVCPRGYILRVTAAYKEGYGEGCVTTHEIENNTDELKFNCEEEVYPTDGTGTCNDYTAH 1130
DB 1023 RVCPRGYILRVTAAYKEGYGEGCVTTHEIENNTDELKFNCEEEVYPTDGTGTCNDYTVN 1082

QY 1131 Q---GTAVCMNSRAGYEDAYEVDVTTASVNYKPTYEEETDVRDNHCEYDRGVVNPPL 1187
DB 1083 QEEYGGAY-TSRNRGYNEAPSV----PADYASVYEEKSVTDGRRENPCFENRGYRDYTPL 1137

QY 1188 PAGYMTKELEYPPETDKVWIEIGETEGKEFIVDSVELLLMEE 1228
DB 1138 PVGYVTKELEYPPETDKVWIEIGETEGTFIVDSVELLLMEE 1178
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 09:24:41 ; Search time 153 Seconds
(without alignments)
4618.040 Million cell updates/sec

Title: US-10-614-524-2
Perfect score: 6479
Sequence: 1 LTSNRKNEINEINALSPAP.....IGETEGKPIVDSVLLMEE 1228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_treml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5924.5	91.4	1227	1 C1BE_BACTU	O85805 bacillus th
2	5745	88.7	1228	1 C1BA_BACTK	P05517 bacillus th
3	5740	88.6	1228	2 Q93T75	Q93T75 bacillus th
4	5737	88.5	1228	2 Q93NM5	Q93nm5 bacillus th
5	5235.5	80.8	1229	1 C1BB_BACTU	Q45739 bacillus th
6	5214.5	80.5	1231	1 C1BD_BACTZ	Q9zaz5 bacillus th
7	5188.5	80.1	1231	2 Q8KNY2	Q8kny2 bacillus th
8	4959.5	76.5	1233	1 C1BC_BACTM	Q45774 bacillus th
9	4284.5	66.1	1215	1 C1CA_BACTM	Q45715 bacillus th
10	3977.5	61.4	1174	1 C1EB_BACTA	Q03745 bacillus th
11	3901.5	60.2	849	2 Q6PYW8	Q6pyw8 bacillus th
12	3901.5	60.2	849	2 AAS93797	Aas93797 bacillus
13	3791	58.5	1176	1 C1CB_BACTG	P56953 bacillus th
14	3579	55.2	1170	1 C1JB_BACTU	Q45716 bacillus th
15	3556	54.9	1167	1 C1JA_BACTU	Q45738 bacillus th
16	3495.5	54.0	1189	1 C1CA_BACTE	P05518 bacillus th
17	3495.5	54.0	1189	2 Q6YNE8	Q6ynb8 bacillus th
18	3495.5	54.0	1189	2 AAM00264	Aam00264 bacillus
19	3495.5	54.0	1189	2 AAL79362	Aal79362 bacillus
20	3495.5	53.8	1189	2 Q91877	Q91877 bacillus th
21	3455.5	53.3	1180	2 Q9S5V8	Q9s5v8 bacillus th
22	3433.5	53.0	1165	1 C1DA_BACTA	P19415 bacillus th
23	3433.5	53.0	1166	1 C1GA_BACTU	Q45746 bacillus th
24	3428	52.9	1179	1 C1AD_BACTA	Q03744 bacillus th
25	3426.5	52.9	1181	1 C1AE_BACTL	Q03748 bacillus th
26	3426	52.9	1171	2 O06894	O06894 bacillus th
27	3426	52.9	1171	2 Q71TW6	Q71tw6 bacillus th
28	3426	52.9	1171	2 AAL50330	Aal50330 bacillus th
29	3425	52.9	1177	2 Q6EIX3	Q6eix3 bacillus th
30	3421	52.8	1171	1 C1SA_BACTX	Q57458 bacillus th
31	3413.5	52.7	1174	2 Q45749	Q45749 bacillus th

ALIGNMENTS

RESULT 1

ID	C1BE_BACTU	STANDARD;	PRT;	1227 AA.
AC	O85805;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Pesticidal crystal protein cryIbE (insecticidal delta-endotoxin CryIbE) (Crystalline entomocidal protoxin) (139 kDa crystal protein).			
GN	Name=cryIbE; Synonyms=cryIB(e), 158C2B;			
OS	Bacillus thuringiensis.			
OG	Plasmid pMYC2383.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1428;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NREL B-19872 / PS158C2;			
RA	Payne J.M., Cummings D.A., Cannon R.J.C., Narva K.E., Steiman S.;			
RT	"Bacillus thuringiensis genes encoding lepidopteran-active toxins.;"			
RL	Patent number US5723758, 03-MAR-1998.			
CC	-!- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut epithelial cells of many lepidopteran larvae.			
CC	-!- DEVELOPMENTAL STAGE: The crystal protein is produced during sporulation and is accumulated both as an inclusion and as part of the spore coat.			
CC	-!- MISCELLANEOUS: Toxic segment of the protein is located in the N-terminus.			
CC	-!- SIMILARITY: Belongs to the delta endotoxin family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AF077326; AAC32850.1; -.			
DR	HSSP; P07130; 1DLC.			
DR	InterPro; IPR001178; Endotoxin.C.			
DR	InterPro; IPR005638; endotoxin.N.			
DR	InterPro; IPR005639; endotoxin.N.			
DR	InterPro; IPR008979; Gal_bind_like.			
DR	Pfam; PF003944; Endotoxin_C; 1.			
DR	Pfam; PF00555; Endotoxin_M; 1.			
DR	Pfam; PF03945; Endotoxin_N; 1.			
KW	Plasmid; Sporulation; Toxin.			
SQ	SEQUENCE 1227 AA; 139084 MW; CBA847BEA0B34CD3 CRC64;			

Query Match 91.4%; Score 5924.5; DB 1; Length 1227;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1128; Conservative 37; Mismatches 60; Indels 5; Gaps 3;

QY 1 LTSNRKNEINEINALSPAPVNSHNHSTQMDLSPDARIEDSLCTAEGNNINPLVSASTVQTGI 60

```
Db 1 MTSNRKNEIINALSIPAVNSHAQMLNSTDARIEDSLCAEGNNDPFFVSASTVQTGI 60
QY 61 NIAGRILGLVGPAGQIASFYSFLVGLWPRGRDQWEIFLEHVEQLINOQITENARNTA 120
Db 61 NIAGRILGLVGPAGQIASFYSFLVGLWPRGRDQWEIFLEHVEQLIROQVIENTRDTA 120
QY 121 LARLOGLGDSFRAYQOQSLEDLWENRDDRARTSVLYTQVIALDELFLNAMPFLAIRNQVEP 180
Db 121 LARLOGLGDSFRAYQOQSLEDLWENRDDRARTSVLYTQVIALDELFLNAMPFLAIRNQVEP 180
QY 181 LLMVYAQAANLHLLLRDASLFGSEFGLTQSEIQYRYERQVEQTRDYSDYCVENYNTGLN 240
Db 181 LLMVYAQAANLHLLLRDASLFGSEFGLTQSEIQYRYERQVEQTRDYSDYCVENYNTGLN 240
QY 241 SLRGTNAASVRYNQFRDLTLGLVLDLVALFPSPYDTRTYPINTSAQLTREVIYTDAGATG 300
Db 241 NLRGTAESWLRYNQFRDLTLGLVLDLVALFPSPYDTRTYPINTSAQLTREVIYTDAGATG 300
QY 301 V--NWASNNWYNNAPSFSAIETAVIRSPHLLDPLEQLTIFPSTSSRWASRTHMYTKRGHT 358
Db 301 APSGFATNFWNNAPSFSAIAEAVIRPPHLLDPPEQLTIFSVLSRWSNTQYNNYWGHR 360
QY 359 IQSRPIGGGLNTSHGSTNTSINPVRLSFRSDRVYMTESVAGVLLGWILYLEPIHGVPTR 418
Db 361 LESRTIRGSLSTSHGNTSINPVLTQTSRDRYRTESFAGINI--LLTPVNGVPPWAR 418
QY 419 FNFNPQNTFERGTANYQPYESFGLQKDSLELPTETTERPNYESYSHRLSHIGLISQ 478
Db 419 FNMENPLNSL-RGSLLYTIGTGTQTLFDSLELPTETTERPNYESYSHRLSNIRLISG 477
QY 479 SRVHVPLYWTHRSADTNTISSDITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
Db 478 NTLRAPVYWTHRSADTNTISSDITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 537
QY 539 GSVLSMGLNFNTSLQRYRVRYAAQOTWLVRYVVGSTTFDQPPSTWSANESLTSQS 598
Db 538 GSVLSMGLNFNTSLQRYRVRYAAQOTWLVRYVVGSTTFDQPPSTWSANESLTSQS 597
QY 599 FRFAEPVVGISAGSQTAGISISNNAQRQTFHFDKIEFIPITATFEAYDLERAQAVNA 658
Db 598 FRFAEPVVGISAGSQTAGISISNNAQRQTFHFDKIEFIPITATFEAYDLERAQAVNA 657
QY 659 LFTWNPRLKTDVTDHIDOVSNLVACLDSEFCLDEKRELLKVKYAKLSDERNLLQD 718
Db 658 LFTWNPRLKTDVTDHIDOVSNLVACLDSEFCLDEKRELLKVKYAKLSDERNLLQD 717
QY 719 PNFTSINKQPDFISTNEQSNFTSIHQSEHGWSGSENIITIQEGNDVPKENYVTLPGTFNE 778
Db 718 PNFTSINKQPDFISTNEQSNFTSIHQSEHGWSGSENIITIQEGNDVPKENYVTLPGTFNE 777
QY 779 CYPYLYQKIGESLKYATRYQLRGYIEDSQDLIEYLIRYNAKHETLDVPGTSEVWPLSV 838
Db 778 CYPYLYQKIGESLKYATRYQLRGYIEDSQDLIEYLIRYNAKHETLDVPGTSEVWPLSV 837
QY 839 ESPIGRCGEFNRCAPHEWNPDDCSCRDGCEKCAHSHHFSLDIDIGCTDLHENLGVWV 898
Db 838 ESPIGRCGEFNRCAPHEWNPDDCSCRDGCEKCAHSHHFSLDIDIGCTDLHENLGVWV 897
QY 899 FKIQSGHARLGNLEPIESKPLLGESALSRVKAERKWRDKREKLQLETKRVVYTEAKEAV 958
Db 898 FKIQSGHARLGNLEPIESKPLLGESALSRVKAERKWRDKREKLQLETKRVVYTEAKEAV 957
QY 959 DALFVDSQYRLQADTNIGIMHAADKLVRIRAYLSLSVIVGVNAEIFEELGRIITA 1018
Db 958 DALFVDSQYRLQADTNIGIMHAADKLVRIRAYLSLSVIVGVNAEIFEELGRIITA 1017
QY 1019 ISLYDARNVYKNGDFNNGLACMNVKGVHDVQQSHRSVLVPEWEAEVSAVRVCPGRGY 1078
Db 1018 ISLYDARNVYKNGDFNNGLACMNVKGVHDVQQSHRSVLVPEWEAEVSAVRVCPGRGY 1077
QY 1079 ILRVYAYKEGYGEGCVTIHIEIENNTDELKPKNCEBEVEVPTDGTGTCNDYTAHQGTACVNS 1138
```

```
1078 ILRVYAYKEGYGEGCVTIHIEIENNTDELKPKNCEBEVEVPTDGTGTCNDYTAHQGTACVNS 1137
1139 RNAGYEDAVEVDVTASVNVKPYTEETTYDVRDNHCEYDRGVVYVPPPLPAGYMTKELEY 1198
1138 RNAGYEDAVEVDVTASVNVKPYTEETTYDVRDNHCEYDRGVVYVPPVAGYMTKELEY 1197
1199 FPETDKVWISIGTEGKFIVDSVELLMEE 1228
1198 FPETDKVWISIGTEGKFIVDSVELLMEE 1227

RESULT 2
C1BA BACTK
ID C1BA BACTK STANDARD; PRT; 1228 AA.
AC P05517; Q45731;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pesticidal crystal protein cryIbA (insecticidal delta-endotoxin
DE CryIb(a)) (crystalline entomocidal protoxin) (140 kDa crystal protein).
GN Names=cryIbA; Synonyms=cryIb(a), cryA4;
OS Bacillus thuringiensis (subsp. kurstaki), and
OS Bacillus thuringiensis (subsp. entomocidus).
OC Bacteria, Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29339, 1436;
RN [1]
RC SEQUENCE FROM N.A.
RC SPECIES=B.t.kurstaki; STRAIN=HD-2;
RX MEDLINE=88203216; PubMed=362680;
RA Brizzard B.L., Whiteley H.R.;
RT "Nucleotide sequence of an additional crystal protein gene cloned from
RT Bacillus thuringiensis subsp. thuringiensis.";
RL Nucleic Acids Res. 16:2723-2723 (1988).
RN [2]
RC SEQUENCE FROM N.A.
RC SPECIES=B.t.entomocidus; STRAIN=HD-110;
RA Soetaert P.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut
CC epithelial cells of insects.
CC -I- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -I- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -I- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06711; CAA29898.1; -
CC EMBL; X95704; CAA65003.1; -
CC FIR; S00873; S00873.
CC HSP; P07130; 1DLC.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; Endotoxin.C.
CC InterPro; IPR005639; Endotoxin.N.
CC InterPro; IPR008979; Gal_bind_Like.
CC Pfam; PF03944; Endotoxin_C; 1.
CC Pfam; PF00555; Endotoxin_M; 1.
CC Pfam; PF03945; Endotoxin_N; 1.
CC Sporulation; Toxin.
FT VARIANT 150 150 Y -> H (in strain HD-110).
SQ SEQUENCE 1228 AA; 139647 MW; C8E3A19FB5D98575 CRC64;

Query Match 88.7%; Score 5745; DB 1; Length 1228;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1098; Conservative 35; Mismatches 90; Indels 10; Gaps 3;
```

```
QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MTSNRKNEIIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVSASTVQTGI 55

QY 61 NIAGRILGLVGPVAGQIASFYSFLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTA 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 NIAGRILGLVGPVAGQIASFYSFLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTA 115

QY 121 LARLQGLGDSFRAYQOOSLEDLENRDDARTSVLTQVIALELDFLNAMPFAIRNQVEVP 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 LARLQGLGDSFRAYQOOSLEDLENRDDARTSVLTQVIALELDFLNAMPFAIRNQVEVP 175

QY 181 LLAMYAQAANLHLLLRDASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVIEWNTGLN 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 LLAMYAQAANLHLLLRDASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVIEWNTGLN 235

QY 241 SLRGTNAASWVRYNQFRDLTLGLVDLVALPSPYDTRTPINTSAQLTREYVTTAIGATG 300
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
236 SLRGTNAASWVRYNQFRDLTLGLVDLVALPSPYDTRTPINTSAQLTREYVTTAIGATG 295

QY 301 VNMAASMWNYNNAFSAIETAVIRSPHLLDFLEQLTIFSSSRWSATRHMTYWRGHTIQ 360
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 VNMAASMWNYNNAFSAIETAVIRSPHLLDFLEQLTIFSSSRWSATRHMTYWRGHTIQ 355

QY 361 SRPIGGGLNTSTHGSTNTSINPVRLSFFSRDYYVTESYAGVLLWGIYLEPIHGVPTRFN 420
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 SRPIGGGLNTSTHGATNTSINPVTLRFPASRDVYRTESYAGVLLWGIYLEPIHGVPTRFN 415

QY 421 FRNPONTERTGANYSQYSPGLOKQDSELPETTERNYESYHRLSHIGLISQSR 480
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 FTNPQINIDRGNTANYSQYSPGLOKQDSELPETTERNYESYHRLSHIGLISQSR 475

QY 481 VHVPPVSWTHSADRTNTISDSITQIPLVKSFNLNGTSVVGPGFTGGDILRNNGS 540
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
476 VNVPPVSWTHSADRTNTIGNRITQIPWVKASELPQGTTVVRGPGFTGGDILRNNTGG 535

QY 541 VLSMLFNNTSLQRYRVRYAASQTMVLRVVGSGSTTFQGGFSTMSANESLTSQSFR 600
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
536 FGPIRVTVNGPLTQRYRGFYASTVDFDFVSRGGTTVNNFRFLRTWNSGDELKYGNFV 595

QY 601 FAEPPVGIASGSGQ-TAGISISNNAQRTQFHPDKIEPIPTATPAEYDLERAQAVNAL 659
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
596 RRAFTPTPTQIQDIIRTSIQGLSGNGEVNIDKIEIIPVTATPAEYDLERAQAVNAL 655

QY 660 FTNTNPRLKTDDVTDYHIDQVSNLVACLSDFCLDEKELKVKYAKRLSDERNLLQDP 719
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
656 FTNTNPRLKTDDVTDYHIDQVSNLVACLSDFCLDEKELKVKYAKRLSDERNLLQDP 715

QY 720 NFTSINKQPDFISTNEQSNFTSIHQSEHGWGWSENITIQEGNDVFKENYVTLPGTFNEC 779
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
716 NFTSINKQPDFISTNEQSNFTSIHQSEHGWGWSENITIQEGNDVFKENYVTLPGTFNEC 775

QY 780 YPTLYQKIGSELKAYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDVPGTESVMPLSVE 839
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
776 YPTLYQKIGSELKAYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDVPGTESVMPLSVE 835

QY 840 SPIGRCEPNRCAPHFENPDLDCSCEGCEKCAHSHHFSIDIDIGCTDLHENLGVWVVF 899
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
836 SPIGRCEPNRCAPHFENPDLDCSCEGCEKCAHSHHFSIDIDIGCTDLHENLGVWVVF 895

QY 900 KIKTOEGHARLGNLEPIBEKPLGALSRVKAERKWRDKEKQLETKRVVYTAKEAVD 959
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
896 KIKTOEGHARLGNLEPIBEKPLGALSRVKAERKWRDKEKQLETKRVVYTAKEAVD 955

QY 960 ALFVDSQVNRLOADTNGIMHAADKLVHRIEAYLSLSVIPGVNAEIFELEGRITAI 1019
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
956 ALFVDSQVNRLOADTNGIMHAADKLVHRIEAYLSLSVIPGVNAEIFELEGRITAI 1015

QY 1020 SLYDARNVVKGDENGLACWNVKGVHDVQOSSHRSVLVPIPEWEAEVSAQVRVCPGCGYI 1079
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1016 SLYDARNVVKGDENGLACWNVKGVHDVQOSSHRSVLVPIPEWEAEVSAQVRVCPGCGYI 1075
```

```
QY 1080 LRVTAKEGYGEGCVTTHIEIENNTDELKFKNCEBEEVYPTDTGTCDNYTAHQGTA---V 1135
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1076 LRVTAKEGYGEGCVTTHIEIENNTDELKFKNCEBEEVYPTDTGTCDNYTAHQGTA 1135

QY 1136 CNSRNAGYEDAYEVDVTTASVNYKPTYEBETVTVRRDNHCEYDRGYVNYPPUPAGYMTKE 1195
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1136 CNSRNAGYEDAYEVDVTTASVNYKPTYEBETVTVRRDNHCEYDRGYVNYPPUPAGYMTKE 1195

QY 1196 LEYPPETDKWIBIGETEGKFIVDSVELLME 1228
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1196 LEYPPETDKWIBIGETEGKFIVDSVELLME 1228

RESULT 3
Q93T75 PRELIMINARY, PRT: 1228 AA.
ID Q93T75
AC Q93T75;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Delta-endotoxin CryIa2.
GN Name=cryIa2;
OS Bacillus thuringiensis (subsp. entomocidus).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1436;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-9;
RA Mat Isa M.N., Abdullah M.A.F., Mahadi N.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF363025; AAK51084.1; -.
DR HSSP; P07130; 1DLC.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; Endotoxin C.
DR InterPro; IPR008979; Gal bind Like.
DR Pfam; PF03944; Endotoxin_C_1.
DR Pfam; PF00555; Endotoxin_M_1.
DR Pfam; PF03945; Endotoxin_N_1.
SQ SEQUENCE 1228 AA; 139620 MW; 3DA2A4DBF59C95C3 CRC64;

Query Match 88.6%; Score 5740; DB 2; Length 1228;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 1097; Conservative 36; Mismatches 90; Indels 10; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MTSNRKNEIIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVSASTVQTGI 55

QY 61 NIAGRILGLVGPVAGQIASFYSFLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTA 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 NIAGRILGLVGPVAGQIASFYSFLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTA 115

QY 121 LARLQGLGDSFRAYQOOSLEDLENRDDARTSVLTQVIALELDFLNAMPFAIRNQVEVP 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 LARLQGLGDSFRAYQOOSLEDLENRDDARTSVLTQVIALELDFLNAMPFAIRNQVEVP 175

QY 181 LLAMYAQAANLHLLLRDASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVIEWNTGLN 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 LLAMYAQAANLHLLLRDASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVIEWNTGLN 235

QY 241 SLRGTNAASWVRYNQFRDLTLGLVDLVALPSPYDTRTPINTSAQLTREYVTTAIGATG 300
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
236 SLRGTNAASWVRYNQFRDLTLGLVDLVALPSPYDTRTPINTSAQLTREYVTTAIGATG 295

QY 301 VNMAASMWNYNNAFSAIETAVIRSPHLLDFLEQLTIFSSSRWSATRHMTYWRGHTIQ 360
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 VNMAASMWNYNNAFSAIETAVIRSPHLLDFLEQLTIFSSSRWSATRHMTYWRGHTIQ 355

QY 361 SRPIGGGLNTSTHGSTNTSINPVRLSFFSRDYYVTESYAGVLLWGIYLEPIHGVPTRFN 420
```

Db 356 SRPIGGGLNTSTHATNTSINPVLRFASRDVYVTESVAGVLLMGVILEPIHGVPVTFN 415
Qy 421 FRNPQNTFERTANYSQPYSPGLQKDSLETLPETTERPNYESYSHRLSHIGLISQSR 480
Db 416 FTNPQNSDRGTANYSQPYSPGLQKDSLETLPETTERPNYESYSHRLSHIGLISQSR 475
Qy 481 VHPVYVSWTHRSADRTNTISDSITQIPLVKSFNLSNGTSVVSQPGTGGDIIRTNVNGS 540
Db 476 VNVVYVSWTHRSADRTNTIGPNRIITQIPMWKASELPQGTTVVRGPGTGGDILRRNTGG 535
Qy 541 VLSMGLFNNTSLQRYRVRYAAASQTMWLVRVTVGGSTTDPQSPSTMANESLTSQSPR 600
Db 536 FGPRIRVTVNGPLTQRYRIGFYASTVDFPFVSRGGTTVANFRFLRTMNSGDELKYGNFV 595
Qy 601 PAEPFVGISAGSQ-TAGISISNNAGRTQHFDFKIEPIPTATFEAAYDLERAQEAVAL 659
Db 596 RRAFTPTFTTQIQIIRTSIQGLSGNGEVYDKIEIIPVTATFEAAYDLERAQEAVAL 655
Qy 660 FTNTNPRRLKTDVTDYHIDQVSNLVACLDFCLDEKRELLKVKYAKRLSDERNLLQDP 719
Db 656 FTNTNPRRLKTDVTDYHIDQVSNLVACLDFCLDEKRELLKVKYAKRLSDERNLLQDP 715
Qy 720 NFTSINKQPDFISTNEQSNFTSIHQSEHGWSGNSNITIQEGNDVFKENYVTLPGTFNEC 779
Db 716 NFTSINKQPDFISTNEQSNFTSIHQSEHGWSGNSNITIQEGNDVFKENYVTLPGTFNEC 775
Qy 780 YPTLYOKIGISELKAATRYQLRGYIYEDSQDLEIYLIRYNAKHETLDVPGTESWPLSVE 839
Db 776 YPTLYOKIGISELKAATRYQLRGYIYEDSQDLEIYLIRYNAKHETLDVPGTESWPLSVE 835
Qy 840 SPIGRCCENRCAPHEFWNPDLDCSCRDGKCAHSHHFLSDIDIGCTDLHENLGVWVVF 899
Db 836 SPIGRCCENRCAPHEFWNPDLDCSCRDGKCAHSHHFLSDIDIGCTDLHENLGVWVVF 895
Qy 900 KIKTQEGHARLGNLEFIEEKPLLGEALSRYKRAEKWRDKREKLQETKEVYTEAKEAVD 959
Db 896 KIKTQEGHARLGNLEFIEEKPLLGEALSRYKRAEKWRDKREKLQETKEVYTEAKEAVD 955
Qy 960 ALFVDSQYNRLADNTNIGMTHAADKLVRIRYAYLSVLSVPIPGVNAIIFBELEGHIIITAI 1019
Db 956 ALFVDSQYDLQADNTNIGMTHAADKLVRIRYAYLSVLSVPIPGVNAIIFBELEGHIIITAI 1015
Qy 1020 SLYDARNVVKNGDFNNGLACWNVKGVHDVQSHRSVLVPIPEAEVSOQAVRVCPCRGYI 1079
Db 1016 SLYDARNVVKNGDFNNGLTCWNVKGVHDVQSHRSVLVPIPEAEVSOQAVRVCPCRGYI 1075
Qy 1080 LRVYAYKEGYGEGCVTHIEIENNTDELKFKXCEEEVYPTDTCTCNDYTAHQGTA---V 1135
Db 1076 LRVYAYKEGYGEGCVTHIEIENNTDELKFKXCEEEVYPTDTCTCNDYTAHQGTA 1135
Qy 1136 CNSRNAGYEDAYEDVDTTASVNYKPYEEETVTVRRDNHCEYDRGVYVPPYPAGYVTK 1195
Db 1136 CNSRNAGYEDAYEDVDTTASVNYKPYEEETVTVRRDNHCEYDRGVYVPPYPAGYVTK 1195
Qy 1196 LEYFPETDKVWIEIGETGKFIIVDSVELLIMEE 1228
Db 1196 LEYFPETDKVWIEIGETGKFIIVDSVELLIMEE 1228

RESULT 4
Q93NM5 ID Q93NM5 PRELIMINARY; PRN; 1228 AA.
AC Q93NM5; DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Cry1Ba.
GN Name=Cry1Ba;
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]

RP SEQUENCE FROM N.A.
RA Zhang J., Song F., Huang D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF368257; AAK63251.1; -.
DR HSSP; P07130; 1DLC.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF03945; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE 1228 AA; 139665 MW; E86D9842341PB439 CRC64;

Query Match 88.5%; Score 5737; DB 2; Length 1228;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 1097; Conservative 35; Mismatches 91; Indels 10; Gaps 3;

Qy 1 LTSNRKNEEIIINALSIPAVSNHSTQMDLSPOARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 MTSNRKNEEIIIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDFVSASTVQTGI 55
Qy 61 NIAGRILGVLPVPPAGQIASFYSFLVGLWPRGRDQWEIPLHVEQLINQITENARNTA 120
Db 56 NIAGRILGVLPVPPAGQIASFYSFLVGLWPRGRDQWEIPLHVEQLINQITENARNTA 115
Qy 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNOEVP 180
Db 116 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNOEVP 175
Qy 181 LLMVYAAANLHLLLRDASLFGSEFGLTSQBIQRYRQVQTRDYSYCVWEYNTGLN 240
Db 176 LLMVYAAANLHLLLRDASLFGSEFGLTSQBIQRYRQVQTRDYSYCVWEYNTGLN 235
Qy 241 SLRGNTAAWVYVNGQRRDLTLGVLDLVALFYSYDTRTPINTSAQLTREVTDAIGATG 300
Db 236 SURGNTAAWVYVNGQRRDLTLGVLDLVALFYSYDTRTPINTSAQLTREVTDAIGATG 295
Qy 301 VNMASNNVYNNAPFSATETAVIRASPHLLDLEOLTTFSTSSRWSSATRHMTYWRGHTIQ 360
Db 296 VNMASNNVYNNAPFSATETAVIRASPHLLDLEOLTTFSTSSRWSSATRHMTYWRGHTIQ 355
Qy 361 SRPIGGGLNTSTHGTNTSINPVLRFASRDVYVTESVAGVLLMGVILEPIHGVPVTFN 420
Db 356 SRPIGGGLNTSTHGTNTSINPVLRFASRDVYVTESVAGVLLMGVILEPIHGVPVTFN 415
Qy 421 FRNPQNTFERTANYSQPYSPGLQKDSLETLPETTERPNYESYSHRLSHIGLISQSR 480
Db 416 FTNPQNSDRGTANYSQPYSPGLQKDSLETLPETTERPNYESYSHRLSHIGLISQSR 475
Qy 481 VHPVYVSWTHRSADRTNTISDSITQIPLVKSFNLSNGTSVVSQPGTGGDIIRTNVNGS 540
Db 476 VNVVYVSWTHRSADRTNTIGPNRIITQIPMWKASELPQGTTVVRGPGTGGDILRRNTGG 535
Qy 541 VLSMGLFNNTSLQRYRVRYAAASQTMWLVRVTVGGSTTDPQSPSTMANESLTSQSPR 600
Db 536 FGPRIRVTVNGPLTQRYRIGFYASTVDFPFVSRGGTTVANFRFLRTMNSGDELKYGNFV 595
Qy 601 PAEPFVGISAGSQ-TAGISISNNAGRTQHFDFKIEPIPTATFEAAYDLERAQEAVAL 659
Db 596 RRAFTPTFTTQIQIIRTSIQGLSGNGEVYDKIEIIPVTATFEAAYDLERAQEAVAL 655
Qy 660 FTNTNPRRLKTDVTDYHIDQVSNLVACLDFCLDEKRELLKVKYAKRLSDERNLLQDP 719
Db 656 FTNTNPRRLKTDVTDYHIDQVSNLVACLDFCLDEKRELLKVKYAKRLSDERNLLQDP 715
Qy 720 NFTSINKQPDFISTNEQSNFTSIHQSEHGWSGNSNITIQEGNDVFKENYVTLPGTFNEC 779
Db 716 NFTSINKQPDFISTNEQSNFTSIHQSEHGWSGNSNITIQEGNDVFKENYVTLPGTFNEC 775

Qy	998	VFKIKTQEGHARLGNLFPIEBKPLLGALSVRKGAEKWRDKREKLOLETKRVTYTEAKA	957
Db	999	VFKIKTQEGYARLGNLFPIEBKPLIGALSVRKGAEKWRDKREKLOLETKRVTYTEAKA	958
Qy	958	VDALFVDSQYNRLQADTNIGMHAADKLVHRIREAYLSELVPIGVNAEIFEELEGRIT	1017
Db	959	VDALFVDSQYDQLQADTNIGMHAADKLVHRIREAYLSELVPIGVNAEIFEELEGHIT	1018
Qy	1018	AISLYDARNVVKNGDFNGLACMNVKHVDVQQSHHRSVLVIPWEAEVSQAVRVCFGRG	1077
Db	1019	AMSLYDARNVVKNGDFNGLTGMNVKHVDVQQSHHRSDLVIPWEAEVSQAVRVCFGRG	1078
Qy	1078	YLIRVTAYKEGYGEGCVTIHEIENNTDELAFKNCBEEBVPTDTGTCDNTYHQGTAVCN	1137
Db	1079	YLIRVTAYKEGYGEGCVTIHEIENNTDELAFKNCBEEBVPTDTGTCDNTYHQGTAACN	1138
Qy	1138	SRNAGYEDAYEVDTTASVNVKPYBEEYTYDVRDNDHCEYDRGVNVPPDPAGVMTKELE	1197
Db	1139	SRNAGYEDAYEVDTTASVNVKPYBEEYTYDVRDNDHCEYDRGVNVPPVPAGVYVTKELE	1198
Qy	1198	YFPETDKVWIEIGBTEGKFIVDSVELLMEE	1228
Db	1199	YFPETDVTWIEIGBTEGKFIVDSVELLMEE	1229

RESULT 6

ID C1BD_BACTZ STANDARD; PRRT; 1231 AA.
 AC Q9ZAS5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Pesticidal crystal protein cryIBD (insecticidal delta-endotoxin
 DE CryIb(d) (crystalline entomocidal protoxin) (140 kDa crystal protein).
 GN Names:cryIBd; Synonyms:cryIB(d), cryIAI, cryEI;
 GN Bacillus thuringiensis (subsp. thuringiensis).
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=52024;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD-525;
 RC MEDLINE=20153386; PubMed=10688690;
 RX Kuo W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-P.;
 RT "Cloning of two new cry genes from *Bacillus thuringiensis* subsp.
 RT *whanensis* strain.";
 RL Curr. Microbiol. 40:227-232(2000).
 CC -!- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut
 CC epithelial cells of lepidopteran larvae. Toxic to *Plutella*
 CC *xylostella*.
 CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -!- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U70726; AAD10292.1; -.
 CC HSP; P02965; 1CIY.
 CC InterPro; IPR001178; Endotoxin.
 CC InterPro; IPR005638; Endotoxin.C.
 CC InterPro; IPR005639; Endotoxin.N.
 CC InterPro; IPR008979; Gal bind_Like.
 CC Pfam; PF03944; Endotoxin_C; 1.
 CC Pfam; PF00555; Endotoxin_M; 1.
 CC Pfam; PF03945; Endotoxin_N; 1.

Db 954 YTEAKETVDALFVDSHYNRLQADTNGIMTHAADRLVHRHIEAYLPPELPIGINAVIFEE 1013
QY 1011 LEGRIITAIISLYDARNVKNQDFNGLACWNVKGHVQOQSHRSVLVTPWEAEVSQAV 1070
Db 1014 LENRISTAFSLYDARNVKNQDFNGLACWNVKGHVQOQSHRSVLVTPWEAEVSQAV 1073
QY 1071 RVCGRGYILRVATYKGGYGGCVTHIEENNTDELKFKNCEBEVYPTDGTCDNYTAH 1130
Db 1074 RVCGRGYILRVATYKGGYGGCVTHIEENNTDELKFKNCEBEVYPTDGTCDNYTAH 1133
QY 1131 OCTAVCNRNAGYDAEYVDVTASVNYKPTVEEYTDVRRDNHCEYDRGVNYPPLPAG 1190
Db 1134 QGTAAACNRNAGYDAEYVDVTASVNYKPTVEEYTDVRRDNHCEYDRGVNYPPLPAG 1193
QY 1191 YMTKELEYFPETDKVWIEIGETEGKFIVDSVELLMEE 1228
Db 1194 YVTKLEYFPETDKVWIEIGETEGKFIVDSVELLMEE 1231

RESULT 7
Q8KNV2 PRELIMINARY; PRT; 1231 AA.
AC Q8KNV2;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE CryIBII.
GN Name=cryIBII;
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Isakov I.A., Isakov Y.B., Rymar' S.E., Yarovoi S.V.;
EL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF138457; AA093496.1; -;
DR HSP; P02965; 1CIY.
DR GO; GO:0005102; P:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin.C.
DR InterPro; IPR005639; endotoxin.N.
DR InterPro; IPR008979; Gal bind like.
DR Pfam; PF03944; Endotoxin_C; 1-
DR Pfam; PF00555; Endotoxin_M; 1-
DR Pfam; PF03945; Endotoxin_N; 1-
SQ SEQUENCE 1231 AA; 139764 MW; C9F2848A9297EA00 CRC64;

Query Match 80.1%; Score 5188.5; DB 2; Length 1231;
Best Local Similarity 79.3%; Pred. No. 9e-318;
Matches 982; Conservative 97; Mismatches 142; Indels 17; Gaps 9;

QY 1 LTSNRKNEIINALSIPAVNSHSTOMDLPDARIEDSLCIAEAGNNINPLVSASTVQGI 60
Db 1 MTSNRKNEIINALSIPVSNPSTQMNLPDARIEDSLCIAEAVNNIDPFVSASTVQGI 60
QY 61 NIAGRILGVLPFAGQATASFYSLVGLWLPGRDOWEFLHVEQLINQITENARNTA 120
Db 61 NIAGRILGVLPFAGQATASFYSLVGLWLPGRDOWEFLHVEQLINQITENARNTA 120
QY 121 LARLOGLGDSFRAYQOQSLMEDLENRDDARTSVLTQVIALELDFPLNAMPFAIRNORVP 180
Db 121 IARLEGLGRGYSYQQALETWLDNENDARSIIILERYVALELDITTAIPFRINEEVP 180
QY 181 LMVYAQANLHLLLRDASLPFGSEFGLTSQEIORYERQVEQTRDYSYCVWYNTGLN 240
Db 181 LMVYAQANLHLLLRDASLPFGSEWMASSDVNYQEQIRYTEYSNHCQWYNTGLN 240
QY 241 SLRGNTNAAWRYNCPFRDLTLGLVLDVALPESYDTRTPYNTSQAQLTREVYTDAGTG 300
Db 241 NLRGNTNAAWRYNCPFRDLTLGLVLDVALPESYDTRTPYNTSQAQLTREVYTDPIGRN 300

QY 301 V--NMASNNWYNNNAPSFSAIETAVIRSPHLLDFLEQLTIPTSTSRWATRHMTYWRGHT 358
Db 301 APSGFASNNWNNNAPSFSAIEAALFRPPLHLLDPPEQLTIYSASSRWSSTQHMTYVWGRH 360
QY 359 IQSRPIGGLNTSTHGST-NTSINPVLRSFPSSRDVYWTESVAGVLLMGIYLEPIHGVPTV 417
Db 361 LNFPIGGLNTSTQGLTNNTSINPVLQFTSRDVRTTESNAGTNI--LFTTPVNGVPMW 418
QY 418 RFNFPNQPNTPERGTANYSQPVESPGLOKDSSETLPPETTERPNYVESYHRLSHIGLIS 477
Db 419 RFNFPNQPNTPERGTANYSQPVESPGLOKDSSETLPPETTERPNYVESYHRLSHIGLII 478
QY 478 QSRVHPVYVSWTHRSADRTNTISSDITQIPLVKSFNLSGTSVSVSGFGTGGDIIRTN- 536
Db 479 GNTLRAPVYVSWTHRSADRTNTIGPNRIQIIPAVKGRFLNG-SVTSFGFGTGGDVVLNR 537
QY 537 VNGSVLSMG-----LNFNNTSLQRYRVRVRAAAGTVMVLRVTVGSGTTFDQGFPTMGAN 591
Db 538 NNGNIQNRGYIEVPIQFTSTS-TRYRVRVRVASVTSIELNVNLGNSSIFTTLPTAASL 596
QY 592 ESLTSQSPPRAEFPVGI-SASGSQTAGISINNAGROTFFHDKTETFIPTATFEAEYDLE 650
Db 597 DNLSGDFGYEINNAFTSATGNIVGARNFSANA---EVIIDRPEFIPVTATFEAEYDLE 653
QY 651 RAQAVNALFTNTNPRRLKTDVTDYHIDQVNLVACLSDSECLDEKRELLKVKYAKRLS 710
Db 654 RAQAVNALFTNTNPRRLKTDVTDYHIDQVNMVACLSDSECLDEKRELLKVKYAKRLS 713
QY 711 DERNLLODPNFTSINKQDPFTSTNEQSNFTSIHQSEHGHWGSGENITTOEGNDVPKENYV 770
Db 714 DERNLLODPNFTSISGQLSFASIDQSNFTSINELSEHGHWGSGENITTOEGNDVPKENYV 773
QY 771 TLPGTENECPYLYOKIGESLKAITYRQYGLGYEDSQDLRLYLIRYNKHEHETLDVPGT 830
Db 774 TLPGTENECPYLYOKIGESLKAITYRQYGLGYEDSQDLRLYLIRYNKHEHETLDVPGT 833
QY 831 ESMVPLSVESPIGRGEPNRCAPHFENWPDLDSCRDGECACAHSHHSHHFTLDIDVGCTDLH 890
Db 834 DSLWPLSVKSPIGRCGEPNRCAPHFENWPDLDSCRDGECACAHSHHSHHFTLDIDVGCTDLH 893
QY 891 ENLGVWVVFVKIQTQEGHARLGNLFIEBKPLLGALSRRVKAERKWRDKREKLQLETQKV 950
Db 894 ENLGVWVVFVKIQTQEGHARLGNLFIEBKPLLGALSRRVKAERKWRDKREKLQLETQKV 953
QY 951 YTEAKEAVDALFVDSQYVNLQADTNGIMTHAADRLVHRHIEAYLPPELPIGINAVIFEE 1010
Db 954 YTEAKEAVDALFVDSHYNRLQADTNGIMTHAADRLVHRHIEAYLPPELPIGINAVIFEE 1013
QY 1011 LEGRIITAIISLYDARNVKNQDFNGLACWNVKGHVQOQSHRSVLVTPWEAEVSQAV 1070
Db 1014 LENRISTAFSLYDARNVKNQDFNGLACWNVKGHVQOQSHRSVLVTPWEAEVSQAV 1073
QY 1071 RVCGRGYILRVATYKGGYGGCVTHIEENNTDELKFKNCEBEVYPTDGTCDNYTAH 1130
Db 1074 RVCGRGYILRVATYKGGYGGCVTHIEENNTDELKFKNCEBEVYPTDGTCDNYTAH 1133
QY 1131 QGTAVCNRNAGYDAEYVDVTASVNYKPTVEEYTDVRRDNHCEYDRGVNYPPLPAG 1190
Db 1134 QGTAAACNRNAGYDAEYVDVTASVNYKPTVEEYTDVRRDNHCEYDRGVNYPPLPAG 1193
QY 1191 YMTKELEYFPETDKVWIEIGETEGKFIVDSVELLMEE 1228
Db 1194 YVTKLEYFPETDKVWIEIGETEGKFIVDSVELLMEE 1231

RESULT 8
C1BC_BACTM
ID C1BC_BACTM STANDARD; PRT; 1233 AA.
AC Q45774;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pesticidal crystal protein cryIbC (Insecticidal delta-endotoxin


```
Db 1 MTSNRKNEIINALSIPAVSNHSAQMNLSLTDARIEDSLCIAEGNNIDPFVSASTVQTGI 60
Qy 61 NIAGRIILGVLPFAGQIASFYSLFVGLWLPGRDQWEIFLEHVQLINQOITENARNTA 120
Db 61 NIAGRIILGVLPFAGQIASFYSLFVGLWLPGRDQWEIFLEHVQLINQOITENARNTA 120
Qy 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLTQVIALELDFLNAMPFAIRNQVVP 180
Db 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLTQVIALELDFLNAMPFAIRNQVVP 180
Qy 181 LMVYAQAANLHLLLRDASLFGSFGSLTQEIQRYYERQVEQTRDYSDYCVWYNTGLN 240
Db 181 LMVYAQAANLHLLLRDASLFGSFGSLTQEIQRYYERQVEQTRDYSDYCVWYNTGLN 240
Qy 241 SLRGTAASWRYNQFRDLTLGLVDLVALPFSYDTRYPINTSAQLTREYVTDIGRTN 300
Db 241 SLRGTAASWRYNQFRDLTLGLVDLVALPFSYDTRYPINTSAQLTREYVTDIGRTN 300
Qy 301 V--NMAAMNWNNAAPSAIETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHT 358
Db 301 V--NMAAMNWNNAAPSAIETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHT 358
Qy 359 IQSRPIGGLNTSTHGTNTSINPVRLSPFSRDVYVTSYAGVLLWGLYLPBIGHVPTVR 418
Db 359 IQSRPIGGLNTSTHGTNTSINPVRLSPFSRDVYVTSYAGVLLWGLYLPBIGHVPTVR 418
Qy 418 NTLRAPVYSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVSGPGFTGDDIIRTNV 537
Db 418 NTLRAPVYSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVSGPGFTGDDIIRTNV 537
Qy 539 GSLSMGLNFNTSLQRYRVRVRYAASQTMVLRVTVGGSTTFDQGFPTMSANESLTSQS 598
Db 539 GSLSMGLNFNTSLQRYRVRVRYAASQTMVLRVTVGGSTTFDQGFPTMSANESLTSQS 598
Qy 598 FRFAEPFVGISASGQTAGISINNAGRTQTHFDKIEFIPITATPEAYDILERAQEAUNA 657
Db 598 FRFAEPFVGISASGQTAGISINNAGRTQTHFDKIEFIPITATPEAYDILERAQEAUNA 657
Qy 659 LFTNTNPRRLKTVDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQD 718
Db 659 LFTNTNPRRLKTVDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQD 718
Qy 719 PNFTSINKQDPFISTNEQSNFTSIHQSEHGWGSENITIQEGNDVPKENYVTLPGTENE 778
Db 719 PNFTSINKQDPFISTNEQSNFTSIHQSEHGWGSENITIQEGNDVPKENYVTLPGTENE 778
Qy 778 CYPTLYQKIGESSELKAYTRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
Db 778 CYPTLYQKIGESSELKAYTRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
Qy 839 ESPIGRGCEPNR 850
Db 839 ESPIGRGCEPNR 849
```

RESULT 12

```
AAS93797
ID AAS93797 PRELIMINARY; PRT; 849 AA.
AC AAS93797;
DT 26-APR-2004 (TReMBLrel. 27, Created)
DT 26-APR-2004 (TReMBLrel. 27, Last sequence update)
DT 26-APR-2004 (TReMBLrel. 27, Last annotation update)
DE Cry1B type crystal protein (Fragment).
OS Bacillus thuringiensis (subsp. kurstaki).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus thuringiensis.
OX NCBI_TaxID=29339;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=K1;
RA Li M.S., Choi J.Y., Roh J.Y., Shim H.J., Boo K.S., Je Y.H.;
RT "cry1B";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY570734; AAS93797.1; -.
FT NON_TER 849
SQ SEQUENCE 849 AA; 95886 MW; PCB98495787CF763 CRC64;

Query Match
Best Local Similarity 60.2%; Score 3901.5; DB 2; Length 849;
Matches 756; Conservative 33; Mismatches 58; Indels 5; Gaps 3;

Qy 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNIDPFVSASTVQTGI 60
Db 1 MTSNRKNEIINALSIPAVSNHSAQMNLSLTDARIEDSLCIAEGNNIDPFVSASTVQTGI 60
Qy 61 NIAGRIILGVLPFAGQIASFYSLFVGLWLPGRDQWEIFLEHVQLINQOITENARNTA 120
Db 61 NIAGRIILGVLPFAGQIASFYSLFVGLWLPGRDQWEIFLEHVQLINQOITENARNTA 120
Qy 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLTQVIALELDFLNAMPFAIRNQVVP 180
Db 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLTQVIALELDFLNAMPFAIRNQVVP 180
Qy 181 LMVYAQAANLHLLLRDASLFGSFGSLTQEIQRYYERQVEQTRDYSDYCVWYNTGLN 240
Db 181 LMVYAQAANLHLLLRDASLFGSFGSLTQEIQRYYERQVEQTRDYSDYCVWYNTGLN 240
Qy 241 SLRGTAASWRYNQFRDLTLGLVDLVALPFSYDTRYPINTSAQLTREYVTDIGRTN 300
Db 241 SLRGTAASWRYNQFRDLTLGLVDLVALPFSYDTRYPINTSAQLTREYVTDIGRTN 300
Qy 301 V--NMAAMNWNNAAPSAIETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHT 358
Db 301 V--NMAAMNWNNAAPSAIETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHT 358
Qy 359 IQSRPIGGLNTSTHGTNTSINPVRLSPFSRDVYVTSYAGVLLWGLYLPBIGHVPTVR 418
Db 359 IQSRPIGGLNTSTHGTNTSINPVRLSPFSRDVYVTSYAGVLLWGLYLPBIGHVPTVR 418
Qy 418 NTLRAPVYSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVSGPGFTGDDIIRTNV 537
Db 418 NTLRAPVYSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVSGPGFTGDDIIRTNV 537
Qy 539 GSLSMGLNFNTSLQRYRVRVRYAASQTMVLRVTVGGSTTFDQGFPTMSANESLTSQS 598
Db 539 GSLSMGLNFNTSLQRYRVRVRYAASQTMVLRVTVGGSTTFDQGFPTMSANESLTSQS 598
Qy 598 FRFAEPFVGISASGQTAGISINNAGRTQTHFDKIEFIPITATPEAYDILERAQEAUNA 658
Db 598 FRFAEPFVGISASGQTAGISINNAGRTQTHFDKIEFIPITATPEAYDILERAQEAUNA 658
Qy 659 LFTNTNPRRLKTVDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQD 718
Db 659 LFTNTNPRRLKTVDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQD 718
Qy 719 PNFTSINKQDPFISTNEQSNFTSIHQSEHGWGSENITIQEGNDVPKENYVTLPGTENE 778
Db 719 PNFTSINKQDPFISTNEQSNFTSIHQSEHGWGSENITIQEGNDVPKENYVTLPGTENE 778
Qy 778 CYPTLYQKIGESSELKAYTRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
Db 778 CYPTLYQKIGESSELKAYTRYQLRGYIEDSDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
Qy 839 ESPIGRGCEPNR 850
Db 839 ESPIGRGCEPNR 849
```

RESULT 13

C1CB_BACTG STANDARD; PRT; 1176 AA.
 AC P56953;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Peptidicidal crystal protein cryICb (Insecticidal delta-endotoxin
 GN Name=cryICb; (Crystalline entomocidal protoxin) (133 kDa crystal protein).
 OS Bacillus thuringiensis (subsp. galleriae).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=29338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD-29.
 RX MEDLINE=93236401; PubMed=8476286;
 RA Kalman S.S.;
 RT "Cloning a novel cryIC-type gene from a strain of Bacillus
 thuringiensis galleriae.";
 RL Appl. Environ. Microbiol. 59:1131-1137(1993)
 CC -|- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut
 epithelial cells of insects. Toxic to Spodoptera exigua and
 Trichoplusia ni.
 CC -|- DEVELOPMENTAL STAGE: The crystal protein is produced during
 sporulation and is accumulated both as an inclusion and as part of
 the spore coat.
 CC -|- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 terminus.
 CC -|- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M97880; -; NOT_ANNOTATED_CDS.
 DR PIR; A48970; A48970.
 DR HSSP; P02965; 1CIY.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin C.
 DR InterPro; IPR005639; endotoxin N.
 DR InterPro; IPR008979; Gal bind like.
 DR Pfam; PF03944; Endotoxin C; 1.
 DR Pfam; PF00555; Endotoxin M; 1.
 DR Pfam; PF03945; Endotoxin N; 1.
 KW Sporulation; Toxin.
 SQ SEQUENCE 1176 AA; 132867 MW; 108233494C2AC018 CRC64;

Query Match 58.5%; Score 3791; DB 1; Length 1176;
 Best Local Similarity 62.3%; Pred. No. 9,6e-230;
 Matches 762; Conservative 125; Mismatches 241; Indels 96; Gaps 16;

QY 40 C1AGNNINLVASSTVQTG---INIAGRILGVL---GVFPAQIASFVSFLVGLWPRG 93
 DB 14 CLSNPEEL--LIDGERISTGSSIDISLSVLQVLSNFEVGGFLVGLLDFVNGIVGP-- 69
 QY 94 RDWEIEFLEHVEQLINQIITENARNTALRLQGLGDSFRAYQOSLEDLENRDDARTSV 153
 DB 70 -SPWDAFLVQLQELINERIAAYARSAAISNLEGLGNFNFIYFAPKEWADPDNPVTR 128
 QY 154 LYTYQYIALEDLFLNAPLFAIRNQEVPLLMVYQAANLHLLRLDASLFGSFGLTQSBI 213
 DB 129 VVDRFRILGDLERDIPSFRIAGFEVPLLSVYQAANLHLLRLDSSIFGARWGLTTIN 188
 QY 214 QRYERQVEQTRDYSDYCVWYNTGLNSLRGTNAASWRYNQFRDLTLGVLDLVALFPS 273
 DB 189 NENYNRLIRHIDEVANHCACTYNGRLNLPKSYQDWITYNRURDLTLTLVLDIAAFPS 248
 QY 274 YDRTFTPTNSAQLTREVYTDAGTGVNMAWMNNYNNNAPSFAIETAVIRSPHLLDFL 333

Db 249 YDNRYPQISVGQLTREIYTDPLITFNPQLQSV---AQLPTFNVMESNAIRTEPHLFDVL 304
 QY 334 EQLTIFSTSRWSATRHMTYWRGHTIQSRPIGGGLNTSTHGSLNTSINPVLRSFFSRDVI 393
 Db 305 NNLTIPTD--WFSVGRNFWGGRHVISNRIGGNITSPYIGREANQEPSPRSFTNGPVF 361
 QY 394 WTESYAGVLLWGLYLBPIHGVPIVR-----FN-----FRNPONTPE---RGT 432
 Db 362 RTLSN-----PTFRPLQQPWAPPFNLRGVEGFESTPLNSFTYRGRGT 405
 QY 433 ANYSQPYESFGLQKDSLETPEPTTERPNYESYSHRLSHIGLISQSRVHV---PVYSW 488
 Db 406 V-----DSLTELPEPDSVPPREGYSHRLCHATFVQSGTPTFLTTPGVFSW 451
 QY 489 THESADRTNTIISDSITQIPLVKSFNLSNGTSSVSGPGFTGGDIIRTNVNGSVLSMLNF 548
 Db 452 THRSATDRNIIPDVINQIPLVKAFLNTSGTSVVRGPGFTGGDIIRTNVNGSVLSMLNF 511
 QY 549 NNTSLQRYRVRYAASQTMWLRLVTVGSTTTDQGFPPSTMSANESITQSOFPAEPVCI 608
 Db 512 SNTTLQRYRVRYAASQTMWVSVTVGGSTTGNQGFPPSTMSANGALTSQSFAPFPPVGI 571
 QY 609 SASGQTAGISISNNAGROTTFHDKIEFIPITATFEAEYDLERAQEAVALNTNPRRL 668
 Db 572 SASGQASISISNNVGRQMFHLDRIEFLPTVSTFEERYDLERAQEAVALNTNQLGL 631
 QY 669 KTDVTDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKELSDERNLLQDPNFTSINKP 728
 Db 632 KTDVTDYHIDQVSNLVECLSDDEFCLDEKRELEKVKYAKELSDERNLLQDRNFRSINGQL 691
 QY 729 DFISTNEQNFSTSIHQSHGWSGSENIITIQEGNDVFKENYVTLPGTNECVPTLYQKI 788
 Db 692 D-----RGWRGSTDITIQQGDDVFKENYVTLPGTDECVPTLYQKI 733
 QY 789 GESELKAYTRYQLRGVIEDSDOLEILRYNAKHETLDPGFTESVWPLSVESPIGRCEP 848
 Db 734 DESKLSYTRYELRYIIEDSQLEILRYNAKHEIVNVPGTGSLWPLSISNSIGPCGP 793
 QY 849 NRCAPHFEMNPDLDCSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWVFKITQEGHA 908
 Db 794 NRCAPHLEWNPDLDCSCRDGEKCAHSHHFSLDIDVGCTDLNEDLGVWVFKITQDQHA 853
 QY 909 RLGNLEFEEKPLLEALSRVKRAEKWRDKREKQLQLETRKRYVTEAKEAVDALFVDSQYN 968
 Db 854 RLGNLEFEEKPLLEALSRVKRAEKWRDKREKLEWETNIVYKEAKESVDALFVNSQYD 913
 QY 969 RLQADTNIGMIHAADKLVRIRREAYLSELSVPGVNAEIPFEELEGRITTAISLYDARNV 1028
 Db 914 RLQADTNIAHIAADKVRIRREAYLPELSVPGVNAGIPEELEGRIFTAYSLYDARNVI 973
 QY 1029 KNGDFNGLACWNVKGHVIV--QQSHRSVLVPEWEAEVQAQVRCVGRGYILRVATYKE 1087
 Db 974 KNGDFNGLLCLWNLKGHVIVVEEQNNHRSVLVPEWEAEVQEVRCVGRGYILRVATYKE 1033
 QY 1088 GYGGCVTIHEIENNTDELKFKNCEEEVPTDTGTCNDYTAHQ---GTAVCSNRNAGYE 1144
 Db 1034 GYGGCVTIHEIENNTDELKFKNCEEEVPTDTGTCNDYTAHQ---GTAVCSNRNAGYE 1092
 QY 1145 DAYEVDTTASVNYKPTVEEETTYDVRDNHCEYDRGVVNYPPPLPAGYMTKELEYFFPDTK 1204
 Db 1093 KSYESNSSVQADVASVYEEKADTDGRDRDNHCSNRNGYGDVTPLPAGYVTKLEYFFPDTK 1152
 QY 1205 VWIEIGTEGKFIIVDSVELLLMEE 1228
 Db 1153 VWIEIGTEGTFIVDSVELLLMEE 1176
 RESULT 14
 C1JB_BACTU
 ID C1JB_BACTU
 AC Q45716;
 DT 30-MAY-2000 (Rel. 39, Created)

30-MAY-2000 (Rel. 39, Last sequence update)
 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Pesticidal crystal protein cryIIb (insecticidal delta-endotoxin
 DE CryIIb) (CryIIb) (CryIIb) (CryIIb) (CryIIb) (CryIIb) (CryIIb) (CryIIb)
 GN Name=cryIIb; Synonyms=cryIIb, cryET1;
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-18973 / EGS092;
 RA von Tersch M.A., Gonzalez J.M. Jr.;
 RT "Bacillus thuringiensis cryET1 toxin gene and protein toxic to
 RT lepidopteran insects.";
 RL Patent number US5356623, 18-OCT-1994.
 CC -!- FUNCTION: Promotes colloid osmotic lysis by binding to the midgut
 CC epithelial cells of many lepidopteran larvae. Toxic to *Plutella*
 CC *xylostella*.
 CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -!- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U31527; AAA98959.1; -
 DR HSP; P02965; 1CIY.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin C.
 DR InterPro; IPR005639; endotoxin N.
 DR InterPro; IPR008979; Gal bind like.
 DR Pfam; PF03944; Endotoxin C; 1.
 DR Pfam; PF00555; Endotoxin M; 1.
 DR Pfam; PF03945; Endotoxin N; 1.
 KW Sporulation; Toxin.
 SQ SEQUENCE 1170 AA; 133553 MW; B053B9619B78DC19 CRC64;
 Query Match 55.2%; Score 3579; DB 1; Length 1170;
 Best Local Similarity 60.5%; Pred. No. 2.2e-216;
 Matches 725; Conservative 125; Mismatches 275; Indels 73; Gaps 17;
 QY 46 NINPL-VSASTVQTGINTAGRLGLVGLVPPAGQIASFSLVGLWPRGRDQWEIFLEHV 104
 DB 31 DIDPLEVMSLLQLLNF-----VPGGFTSLGFDKINGALRP---SDWELFLAQI 79
 QY 105 EQLNQITENARTALRGLGDSFRAYQOQSLDMLNDRDARTSRVLYQVIAELD 164
 DB 80 EQLIDQTEATVRAKIAIELEGISGRFQLYVEAPKEWETPDNTAARSRTVERFIIDAQ 139
 QY 165 FLNAMPLAIRNOEVPILMVYQAQANHLILLRDLASLFGSBEFLTSQEIQRYEROVQT 224
 DB 140 IEANIPSPRIQFGEVPLLSVYQAQANHLALLRDSVIFGERWGLTNTVNDIYNQVKRI 199
 QY 225 RDYSDYCVETWNTGLNSRGNTAASWRYNQRDLTLGLDLVALPPSYDTRTYPIINTS 284
 DB 200 HEYSDHCVDYTKELERLGFSTRQWKYNGFRRELTUTLVDIVAVFPNYDGLKYPQTK 259
 QY 285 AQLTREYVTDAGATGVNMAWNNMNNAPSFAIET-AVIRSPHLDFLQLTIFSTSS 343
 DB 260 SELTREIYTSVP--SEYYGAINNYQN-----GIOTERQIRQPHLMDFFNTMTWTSYN 312
 QY 344 R-----WSATRMTYWRGTIQSRPIGGGLNTSTHGSTWTSINPRLSPFSDVDVWTSYA 399
 DB 313 RREYVWGLEMTAYETGFA-----GPQVSPFLAGTRGDAAPPFNVRNSDGIY----- 360

QY 400 GVLLWGIYLEIHGVPTVRFNFRNPONTFERGTANYSQP-----YESPGLQKDSSELEPP 455
 DB 361 RILSAPPYSAPFLG--TSVLGSRGEFMF--ALNNISPPPSARYNPG--TVDSLVSIPP 414
 QY 456 ETTERPNVESYSHRLSHLIGLSQSRVHVVPVSWTHRSADRTNTTSSDITQIPLVKSNNL 515
 DB 415 QDNSVPPHRRGSSHLSSHVTWENSS---PIFHWTHRSATTTNINSDVITQIPMKAYNNL 470
 QY 516 NSGTSVSVSGPGCTGDIIRTNVNGSVLSMGLNFNTSLQRYRVVRVRYAASQTMVLRTVVG 575
 DB 471 HAGATVVRGPGTGTGDIIRTSNGMVTLRVDASAVNRQYRIIRFYATNFFYFVVRRG 530
 QY 576 GSTTFDQGFPTSGANESLTSQSFRFAEPFVGISASGSQT-AGISISNNAQRQTFFPKI 634
 DB 531 NLGVNGREIMKTMSTGRELKASAVLGEFTTPFPFNQVPLQIEIQSLSPGGEVYLDKI 590
 QY 635 EFIPITATFEAYDLERAQBAVNAALPTNPNRRLKTDVTDYHIDQVSNLVACLSEPCLD 694
 DB 591 EFIPADTTFEAYDLERAQAVNALFTSTNORGLKTDVTDYHIDQVSNLVCLSEPCLD 650
 QY 695 EKRELEKVKYAKELSDERNLLQDPNFTSINKQDPFISTNQSNFTSIHQSEHCWQSE 754
 DB 651 EKRELEKVKYAKELSDERNLLQDPNFTSINKQDPFISTNQSNFTSIHQSEHCWQSE 692
 QY 755 NITIQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKAATRYQLRGYIEDSODLEIY 814
 DB 693 DITIQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKAATRYQLRGYIEDSODLEIY 752
 QY 815 LIRYNAXHETLDVPGTSGVWPLSVESPIGRGSPNRCAPHFEMWNPDDDCSCRDGEKCAH 874
 DB 753 LIRYNAXHETLDVPGTSGVWPLSVESPIGRGSPNRCAPHFEMWNPDDDCSCRDGEKCAH 812
 QY 875 SHHPSLDIDIGCTDLHENLGVVVFVKITQGHARLGNLEFIEKPLIGELSRVRAEK 934
 DB 813 SHHPSLDIDIGCTDLHENLGVVVFVKITQGHARLGNLEFIEKPLIGELSRVRAEK 872
 QY 935 KWRDKREKLQLETXRVVTEAKEAVDALFVDSQYNRLQADTNIGMHAADKLVRIRERAYL 994
 DB 873 KWRDKREKLQLETXRVVTEAKEAVDALFVDSQYNRLQADTNIGMHAADKLVRIRERAYL 932
 QY 995 SELSVIPGVNAIPEELEEGRIITAIISLYDARNVYKNGDFNGLACWNVKGVHDVQOSHHR 1054
 DB 933 PELFPIPGINAIPEELEENRISTAFLEYARNVINGDFNGLACWNVKGVHDVQOSHHR 992
 QY 1055 SVLVIPSEWAEVSQAVRVCPCRGVILRVATYKEGCGCVTIHEIENNTDELKFCNCEE 1114
 DB 993 SVLVIPSEWAEVSQAVRVCPCRGVILRVATYKEGCGCVTIHEIENNTDELKFCNCEE 1052
 QY 1115 EYVPTDGTGNDYTAHQCTA-----VCNSRNAGYEDAYEDVDTASVNVKPYEEETTDVR 1170
 DB 1053 GDYSNDGTGNDYTAHQCTA-----VCNSRNAGYEDAYEDVDTASVNVKPYEEETTDVR 1112
 QY 1171 RDNHCEYDRGVYVNPPLPAGYMTKELEYFPETDKVWIEIGESTGEKFIQVDSVELLMEE 1228
 DB 1113 EDNHCEYDRGVYVNPPLPAGYMTKELEYFPETDKVWIEIGESTGEKFIQVDSVELLMEE 1170
 RESULT 15
 C1JA BACTU
 ID C1JA BACTU STANDARD; PRT; 1167 AA.
 AC Q45738;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE Pesticidal crystal protein cryIIa (insecticidal delta-endotoxin
 DE CryIIa) (CryIIa) (CryIIa) (CryIIa) (CryIIa) (CryIIa) (CryIIa) (CryIIa)
 GN Name=cryIIa; Synonyms=cryIIa, cryET4;
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-21110 / EGS847;

